



QY 302 TCCGCAATGAATTGTGCACATGGCGTCTAGTCGTGCAGCTTTGGAAACGTAAGA 361  
| | | | | | | | | | | | | | | | | | | | | |  
Db 305 CACCAGAGAGATTATTTGTTTGGGTTTAGTTTTAATAAAAATAAGCATTAATAAC 364  
| | | | | | | | | | | | | | | | | | | | | |  
QY 362 AAAAGATTCGATCCTTAAACGATTTAGAGAAAAAATATAGATCTATTGGATACAGTTTAC 421  
| | | | | | | | | | | | | | | | | | | | | |  
Db 365 AAAAGATTAAAGAAATCAATATTAATGAATGABAAAAAATAACACTTGAATATCGTTCCG 424  
| | | | | | | | | | | | | | | | | | | | | |  
QY 422 CGTTCACTGATATTTGTACAACCGAGCGGCCGAAACCTGACATAGAACCAAGACAG 481  
| | | | | | | | | | | | | | | | | | | | | |  
Db 425 CTTCACATTTATGCTGTCATGATGAAACAGGACATACCGTACCATAGAACCTCAAAATG 484  
| | | | | | | | | | | | | | | | | | | | | |  
QY 482 ACGGCGCTCAAAGTCTATGATATATCAACCTGGGATGATGACGAACAGCCCCGACTTAAAT 541  
| | | | | | | | | | | | | | | | | | | | | |  
Db 485 GCTATTATTAATGTTAAAGTATATTAATGTTCACTTAACAAATGAACTTAATTAGATT 544  
| | | | | | | | | | | | | | | | | | | | | |  
QY 542 GGCAATTAACCAATCTGCAGCAATTAACAGAAATGACCGAAGCAATTAGAGCAAAAG 601  
| | | | | | | | | | | | | | | | | | | | | |  
Db 545 GGCATCTATCTAACTTAAGAAATTAACGCTTTTAAACGCCACAGAAATCACCAATCAAT 604  
| | | | | | | | | | | | | | | | | | | | | |  
QY 602 AGATGGGGGAGTTAGCCCTTTCTGCTTTTGGCCAAAGGCTTAGAACTGTGTGCTGCCGG 661  
| | | | | | | | | | | | | | | | | | | | | |  
Db 605 TAATRGTTAAAGTCTAGTAGATCAATGGGCTGTGAGACAGAAACAAATGGCTTAACGG 664  
| | | | | | | | | | | | | | | | | | | | | |  
QY 662 GGGATTATACACCGGCTTCCCAGTTTGTCAAGCTGTTATTTGAAAGCAATCTAGAAC 721  
| | | | | | | | | | | | | | | | | | | | | |  
Db 665 GTGGTTATACGTCACACAGATCGTTTATACGGCTACATATTTAAGACACCAATACGCT 724  
| | | | | | | | | | | | | | | | | | | | | |  
QY 722 CGGCGCGCCGATGAACGAAGGTGAACAGCCGCTTTCAATTTTGGCAATATAGACA 781  
| | | | | | | | | | | | | | | | | | | | | |  
Db 725 GTTCCCATTAATGAATGAATGAATTTAATGAATTTTAAAGTTTCTAGAAATCAGTACAGA 784  
| | | | | | | | | | | | | | | | | | | | | |  
QY 782 TACCAAGAGGCGCAGTGAATACGGAAGAAAGCAATTCATTAATCGCAATTAATCTCCG 841  
| | | | | | | | | | | | | | | | | | | | | |  
Db 785 TCCTCAAGAGTGCATTTAT--CGATGCCAATRAAATACATTAACACAAATTCATTAAG 841  
| | | | | | | | | | | | | | | | | | | | | |  
QY 842 TGATGTGCAACGAACGAACTGAACTACTATTTCACACACTATGACAAATCGGCAATCCAA 901  
| | | | | | | | | | | | | | | | | | | | | |  
Db 842 TGATGAAAGTAAAGAAAGAAAGTATTAATTAAGCTTAATTTGCAATCAAAATTTTCA 901  
| | | | | | | | | | | | | | | | | | | | | |  
QY 902 AAGTAATTTAATTCATGA 920  
| | | | | | | | | | | | | | | | | | | | | |  
Db 902 AAATTAATTAATCTAGAGA 920  
| | | | | | | | | | | | | | | | | | | | | |

RESULT 2  
US-09-134-001C-2421  
Sequence 2421, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2421  
LENGTH: 996  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2421

Query Match	16.0%;	Score 159;	DB 3;	Length 996;
Best Local Similarity	49.8%;	Pred. No. 1,4e-39;		
Matches 458;	Conservative	0;	Mismatches 455;	Indels 6;
				Gaps 2;
QY	2	ATATGTGCACAAGCTTTACATTGGAAACTGCTGACCGTAAACATGTATTGGCAAGACAA	61	

Db	14	ATATGTGACGCCATTTCTTATATACAAACAAACGTTACATTATTTAGCTAGAACAA	73
Qy	62	TGAAATTTGCATTTCACTGCGGGACAGAGGTGATCTCTATCCGCGCTTACAGTGGAC	121
Db	74	TGACCTTTGGCATTTGAATTTATGGTATCCCAACCATTTGTTCCAGCCATTATCACTACC	133
Qy	122	ACAGTGAAGCTGACGGAAGGGCCCATCAGACACAGTACGCGTTATCGGTATAGGGAGAA	181
Db	134	AAATTGATCTAGATTCAGACATGGGCTCTT---GAATATGGTTTGTGGACAAATTTAA	190
Qy	182	AACTTGAAATATATATTTATTTGCGGACGGCATTTATGAAGCGGTTATCTTGCGCGCC	241
Db	191	AAGTAGACGTTATAGATTTTGTGTATGGTATTAACGAAAAAAGTTTACTATTTGGAAC	250
Qy	242	TTTATTTTCCGGGCTATGCGGAGTACGAAAAACGATACGGGAGATACCTTCACATTC	301
Db	251	ATTACTTCACTGGTGAAGCCTCATACAGTACCCATTAACGTTATGGTTATTTTAACTTAG	310
Qy	302	TCCCGCATGATTTGTGACATGGGGCTGTCAAGTCGTCACTCTTGGAAGCGTAAAG	361
Db	311	CACCTGAGGATTTATGTGTGGGTTTAGGTTTAAATPAAAGTATAGCGAATTAAC	370
Qy	362	AAAGATTCGATCTTTAACGATTTGAGAGAAAAATTTGATCTATGTAGTACAGTTTAC	421
Db	371	AAAGGTTAAGAAAAATCAATTTATGATGATGAAAAAATAACGACTTTGAAATATCGTTC	430
Qy	422	CGCTTCACTGATATTTGTCAAGCCGGACGGCGCGAAACCTGACGATAGAACAGACGAC	481
Db	431	CTTTACATTTCACTGGTCACTGATGAAACAGGACATACGTAAGCCATAGAACCTCAAAATG	490
Qy	482	ACGGCTCAAAAGTCTATGATTAATCACTGGGTGTATGACGAAACAGCCCGACTTATAT	541
Db	491	GCTTATATATGTATPAAAGATATATATATATGTTCACTTAACAATGAACTTAATTAATAT	550
Qy	542	GGCATGTATCAAACTGACGACATATACAGGAATCAGACCGAAGCAATTAGAGAGCAAG	601
Db	551	GGCATCTATCTTAACTTAAAGAAATTAACGCTTTTAAACGCAACAGAAATACCAATCAAT	610
Qy	602	AGATGGCGGATTAAGCCCTTCTCTGCTTTTGGCCAGAGGCTTAGAACTGTGTCTGCGG	661
Db	611	TAAATGTAAAGTGTAGTAAAGATCAATAGGGGTGGAACAGAACAAATGGCTTACCGG	670
Qy	662	GGGATTAATACACCGCTTCCGGGTTTGTCAAGCTGTTTATTTGAAAGAACATCTAGAGC	721
Db	671	GTTGTTATACGTCAAACAGATCGTTTATTAACGGCTACATATTTAAGACACCAACTCGCT	730
Qy	722	CGGCGCCGATGAAACGAAAGGTAAACAGCCGCTTTCAAATTTGGCAAAATATACGA	781
Db	731	GTTCCCATTAATGMAATAGAAATTTAAAGAAATGTTTAAAGTTCTAGAAATCAGTACGTA	790
Qy	782	TACCAAAAGGCGCAGTGAATTAACGAGAAAGACGAATTCATTTATCGCAATATACTTCG	841
Db	791	TCCCTCAAGGTGCACTTAT---CGATGCCAATTAATATCATTTACACAAATATCAATTAG	847
Qy	842	TGATGTGCAACGAAACGTGAAACTATATTTTCCACACATATGACATCGGCAATCCAA	901
Db	848	TGATGGAAGTAAAGAAAGATTTATTTATTAACCTTACTTGTAGCAATCAAAATTTTCA	907
Qy	902	AAGTAAATTTATTTATGTA 920	
Db	908	AAATTAATTAACGTGAAG 926	

```

RESULT 3
US-09-710-279-3345/C
; Sequence 3345, Application US/09710279
; Patent No. 6703492
;
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS48US
; CURRENT APPLICATION NUMBER: US/09/710.279
; CURRENT FILING DATE: 2000-11-09

```

PRIOR APPLICATION NUMBER: 60/164,258  
 PRIOR FILING DATE: 1999-11-09  
 NUMBER OF SEQ ID NOS: 4472  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3345  
 LENGTH: 2975  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 OTHER INFORMATION: nucleic acid sequence  
 US-09-710-279-3345

Query Match 16.0%; Score 159; DB 3; Length 2975;  
 Best Local Similarity 49.8%; Pred. No. 2.5e-39;  
 Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTCACAACTCTTACATTTGGAACTGCTGACCGCTAAACATGTATTAGCAAGAACAA 61  
 DB 1755 ATATGTCACAACTCTTACATTTGGAACTGCTGACCGCTAAACATGTATTAGCAAGAACAA 1696  
 QY 62 TGGATTTTGCATTTAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 121  
 DB 1695 TGGATTTTGCATTTAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 1636  
 QY 122 ACAGTGAAGCTGACGGAAGGGCCCATCAGACAGTACCGGTTATCGGTATGCGGAGAA 181  
 DB 1635 AATTGATCTAGATTCAGATCAGATCGCTT---GAATATGTTTGTGGAAACAAATTTAA 1579  
 QY 182 AACTTGGAAATATATATTATTTGCGAGCGCATTTATGAAACGGTTATCTTGTGCGCGC 241  
 DB 1578 AAGTAGAGCTTATGATTTGTGATGTATTAACGAAAGGTTTATGATTTTCAAC 1519  
 QY 242 TTATTTTCCGGCTATGCGAGTACGAAACAAACGATACGGAAGATACCGTTACATTTG 301  
 DB 1518 ATTACTTCACTGTGTAGAGCTCATACGATCCCATTAACCTTATGTATTATTTACCTTAA 1459  
 QY 302 TCCCGCATAGTTTGTGACATGGGTGCTGCTGCTGCTGCTGCTTGGAAACGATAAAG 361  
 DB 1458 CACCTGAGAGTTTATGTTTGGGTTTATGTTTATTAAGTATTAAGCAATTTAAAC 1399  
 QY 362 AAAAGTTGATTTTAAAGATTTGTAGAGAAATTAATGATCTATTTGATACGTTTAC 421  
 DB 1398 AAAAGTTGATTTTAAAGATTTGTAGAGAAATTAATGATCTATTTGATACGTTTAC 1339  
 QY 422 CGCTTCACTGATATTTGTGACAGCGGCGGAAACCTGACGATGAACCAAGACAG 481  
 DB 1338 CTTTACATTTCACTGATATTTGTGACAGCGGCGGAAACCTGACGATGAACCAAGACAG 1279  
 QY 482 ACGGCTCAAACTCTATGATTAATCAACTGTGTCTATGACGAAACAGCCCGACTTATAT 541  
 DB 1278 GCTTATTAATTAAGTAAAGATTAATTAATGTTTCACTTAACTTAAGCAATTAATTAAGT 1219  
 QY 542 GGCATGTAACCAATCTGACGATATTAAGAGATTCAGACGGAACCAATTAAGAGCAAG 601  
 DB 1218 GGCATGTAACCAATCTGACGATATTAAGAGATTCAGACGGAACCAATTAAGAGCAAG 1159  
 QY 602 AAGTGGCGGATTAAGCCCTTTTGTCTTTTGGCCAAAGCTTAAAGATGTTGTGCTGCGG 661  
 DB 1158 TATATGATTAAGATGCTAGTATGATCAATGGGCTGTGAGAGCAAGAAATGCTTACCG 1099  
 QY 662 GGGATTAATACCGCTTCCCGGTTTGTGACAGCTGTTTATTTGAAAGAACATCTAGAGC 721  
 DB 1098 GTGGTATATAGCTACAGATCGTTTATAGCGCTACATATTTAAAGACCAACCTAGCGCT 1039  
 QY 722 CGGCGGCGCATTAAGAAAGGTGTAACAGCGCTTTTCAAAATTTTGGCAAAATATACGA 781  
 DB 1038 GTTCCATATATAGAAAGTAAATTTATGATTTGTTTAAAGTTCTAGAAATCAGTACGA 979  
 QY 782 TACCAAGGCGCATGATTAAGAGAAAGCAAGAAATTCATTTAGCAATATATCTTCCG 841  
 DB 978 TCCCTCAAGGTGCAATTAAT---CGATGCCAATTAATATCTTACACAAATATCATTTAG 922

QY 842 TGATGTCAACGAAGTGAAGAACTACTATTTCACCATATATGCAATCGGCAATCCAA 901  
 DB 921 TGATGTCAACGAAGTGAAGAACTACTATTTCACCATATATGCAATCGGCAATCCAA 862  
 QY 902 AACTAATTAATTTCACTGA 920  
 DB 861 AATTAATTAATTTCACTGA 843

RESULT 4  
 US-09-710-279-4184  
 Sequence 4184, Application US/09710279  
 Patent No. 6703492  
 GENERAL INFORMATION:  
 APPLICANT: KIMBERLY, WILLIAM JOHN  
 TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 FILE REFERENCE: P0348005  
 CURRENT APPLICATION NUMBER: US/09/710,279  
 PRIOR FILING DATE: 2000-11-09  
 PRIOR APPLICATION NUMBER: 60/164,258  
 NUMBER OF SEQ ID NOS: 4472  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 4184  
 LENGTH: 4114  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 OTHER INFORMATION: nucleic acid sequence  
 US-09-710-279-4184

Query Match 16.0%; Score 159; DB 3; Length 4114;  
 Best Local Similarity 49.8%; Pred. No. 2.9e-39;  
 Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTCACAACTCTTACATTTGGAACTGCTGACCGCTAAACATGTATTAGCAAGAACAA 61  
 DB 1626 ATATGTCACAACTCTTACATTTGGAACTGCTGACCGCTAAACATGTATTAGCAAGAACAA 1685  
 QY 62 TGGATTTTGCATTTAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 121  
 DB 1686 TGGATTTTGCATTTAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 1745  
 QY 122 ACAGTGAAGCTGACGGAAGGGCCCATCAGACAGTACCGGTTATCGGTATGCGGAGAA 181  
 DB 1746 AATTGATCTAGATTCAGATCAGATCGCTT---GAATATGTTTGTGGAAACAAATTTAA 1802  
 QY 182 AACTGGAATATATATTATTTGCCGACGCGATTAAAGAAAGGTTTATCTTGTGCGGCGC 241  
 DB 1803 AAGTAGAGCTTATGATTTGTGATGTATTAACGAAAGGTTTATGCTATTTCAAC 1862  
 QY 242 TTTATTTTCCGGCTATGCGAGTACGAAACCAATTAAGAGATTAACGTTCACTG 301  
 DB 1863 ATTACTTCACTGTGAGAGCTCATACGATTAAGGTTATGTTTAACTTTAACTTTAG 1922  
 QY 302 TCCCGCATAGTTTGTGACATGGGTGCTGCTGCTGCTGCTGCTTGGAAACGATAAAG 361  
 DB 1923 CACCTGAGAGTTTATGTTTGGGTTTATGTTTATTAAGATTAAGCAATTTAAAC 1982  
 QY 362 AAAAGTTGATTTTAAAGATTTGTGACAGCTGTTTATTTGAAAGAACATCTAGAGC 421  
 DB 1983 AAAAGTTGATTTTAAAGATTTGTGACAGCTGTTTATTTGAAAGAACATCTAGAGC 2042  
 QY 422 CGCTTCACTGATATTTGTGACAGCGGCGGAAACCTGACGATGAACCAAGACAG 481  
 DB 2043 CTTTACATTTCACTGATATTTGTGACAGCGGCGGAAACCTGACGATGAACCAAGACAG 2102  
 QY 482 ACGGCTCAAACTCTATGATTAATCAACTGTGTCTATGACGAAACAGCCCGACTTATAT 541  
 DB 2103 GCTTATTAATTAAGTAAAGTAAATTTATGTTTCACTTAAAGATTAAGCAATTAATTTAG 2162  
 QY 542 GGCATGTAACCAATCTGACGATATTAAGAGATTCAGACGGAACCAATTAAGAGCAAG 601





OY 964 GAAGAGTATTCATGACCTTAATTA 989  
DB 2731 CATATCACTATTAGAAAGTTGAATGA 2706

## RESULT 6

US-08-781-986A-444/C  
Sequence 444, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 444:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11466 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-444

Query Match 14.7%; Score 146; DB 3; Length 11466;  
Best Local Similarity 47.5%; Pred. No. 6.3e-35;  
Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

OY 4 ATGTGCAAGCTTTCATTTGAACTGCTGACCGTAAACATGTATTAGCAAGAACATG 63  
DB 3688 ATGTGCAAGCTTTCATTTGAACTGCTGACCGTAAACATGTATTAGCAAGAACATG 3629  
OY 64 GATTTCGATTCAGCTGGGAGACAGAGGTATCTTATCCGCGCGTTACAGCTGAGAC 123  
DB 3628 GATTTCGATTCAGCTGGGAGACAGAGGTATCTTATCCGCGCGTTACAGCTGAGAC 3569  
OY 124 AGTGAAGCTGACGGAAGGCGCCATCAGACAGAGTACCGCTTATTCGCTATGGGAGAAAA 183  
DB 3568 TCTTGCACT--GCAACGACAGGCGCAACGCAATATGCTTTATTTGGCAGAGAAACGAT 3512  
OY 184 CTGGAATATATATTTATTTGCGGACGCGCATTAATGAAGCGTTATCTTGCGCGCGCTT 243  
DB 3511 ATGGAAGTTTATTTATTTATGCTGATGCTGTTATGTAATGATGCGCTTGCAATTCACACA 3452  
OY 244 TATTTTCGGGCTATGCGAGTACGAAAAACGATACGGAAGATACCGTTCACTTGC 303  
DB 3451 TATTTTCGGGCTATGCGAGTACGAAAAACGATACGGAAGATACCGTTCACTTGC 3392  
OY 304 CGCATGAGTTTGTGACATGGGCTGCTCAGTCTGTCAGCTTTGGAAGCGTAAAGAA 363

DB 3391 CAAATGAAATTTGACATGATTTTGGATATACAAAGCATTTGAAGATATGAAACA 3332  
OY 364 AAGATGATCTTTAAGATTTGAGAGAAAAATTAATCATGATGATGATGATTTACCG 423  
DB 3331 CAGACATCCCAATATACATGTTGTAGCTGTATTTAATGATCATGCGGTGAAGTCCGCA 3272  
OY 424 CTTCACGTATATTTGTCAGACCGGACGCGCGCAAACTGATGATGAAACCAAGACAC 483  
DB 3271 TTGATATATCATGTTTCCGATGCACTGACATACATGCAAGTTTCATTTAAGAGGGT 3212  
OY 484 GGCTCAATCTATATGATTAATCAACCTGCTGTATGACGAAACAGCCCGCATTTATATG 543  
DB 3211 GAAGTGTATTAAGAGTAACTCTATGCTGTCTTAACAATCATCAAGCTTAAATGG 3152  
OY 544 CATGTAACTATGCAATGCAATATACAGATGACGACGAAACCAATTTAGAGCAAAAG 603  
DB 3151 CATTTATGTAATTTAAGACATATATATATTTCTCTTATCCAGCAACGAAATTTA 3092  
OY 604 ATGGCGGATTAAGCCCTTTCTGCTTTGGCCAAAGCTTGAACCTGTGTCTGCGCGG 663  
DB 3091 TTGGAAGGTATTAAGATTAACCTTTAGCAATGAAAGACGATACATTTGATTCAGGT 3032  
OY 664 GATTATACACGCGCTTCCGCTTTGTCAGAGCTGTTATTTGAAAGAACTTAGACCG 723  
DB 3031 GAATTAATTTCACTGACGCGCTTTGTAAGATGCAATTTAAGAACCAATTCCTCAA 2972  
OY 724 GCGGCGATGAAAGAAAGGTATACAGCGCTTTTCAATTTTGGCAATATGACGATA 783  
DB 2971 AACATATATTAAGAAATGATTTAATGATGATTTATTTATTAATGATGCGGTAAATTA 2912  
OY 784 CCAAGGCGCAGTGTATTAACGAAAGAAAGCAAAATCATTTATGCAATATTAATCTCCGT 843  
DB 2911 CCGATTTGAATTTGATGCGCGATGATGCTGACATACATCATATGATGATGACGCTA 2852  
OY 844 ATGTGCAAGCAATCTGGAATCTATTTTCCACCATATGACATATGCGCAATTCGAAAA 903  
DB 2851 ATAAATTTAATCTCAAGAAAGTATATATTAATGATTTATGACGATGATTAATGATGCA 2792  
OY 904 GTAATTTATTTATGATGAACTTTGCTGTTGAGCGTAAAGTGTTCGCGTAAAGCA 963  
DB 2791 TTTAAGCTCAGATATATTTAATTAAGAAAAATATGACATTTTAAAGCTGAGAG 2732  
OY 964 GAAGAGATATTCATGAGCTTAATTA 989  
DB 2731 CATATCACTATTAGAAAGTTGAATGA 2706

## RESULT 7

US-09-134-000C-838  
Sequence 838, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stramm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 838  
LENGTH: 987  
TYPE: DNA  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-838

Query Match 9.6%; Score 95.4; DB 3; Length 987;  
Best Local Similarity 45.8%; Pred. No. 1.6e-19;  
Matches 407; Conservative 0; Mismatches 476; Indels 6; Gaps 2;

OY 41 AACATGATTAACAGAAACATGATTTTTCAGCTGGGAGACAGAGGTGATTTCTCT 100

DB 44 ATCAATTAATTTGAAAGAAATTTGATTAATGAAATTTCTTAATGAGGCTGTTACTATTA 103  
 QY 101 ATCCGCGCCCTAGAGCTGGAACAGTGAAGCGGAGGCGCATGACAGCTAG 160  
 DB 104 CGCCGAGAAATTAATAGTTTTCATTCGGAAGTTGGAATTTAGATCATTTTGA 163  
 QY 161 CGTTATGCGATAGGAGGAGAAATTTGAAATATATATTTTCCGACGCGCATTAATGAA 220  
 DB 164 TAATTTGAAATGCTGCTGGGATTTGCTGATTAATCCGCTTAATATGCAATTAATGAAA 223  
 QY 221 GCGGTTATCTTTCGCGGCTTATTTTCCGCTTAATGCGAGTACGAAATTAATGATC 280  
 DB 224 AAGGATTAAGAAATGCTGATTAATTTTCAAGCTATGCAATTAATTAATTAATTAAT 280  
 QY 281 GGAAGATACCGTTACATGTCGCGCATGATTTGATGATGAGGCTGCTAGTCTGTC 340  
 DB 281 AAGAAAGAAATTAATGTTTCTCCATTTGATTAATTTCTTGGGATTTGGCCCATGCT 340  
 QY 341 AGCTTTGGAAGACGTTAAAGAAATTCGATCTTTAAAGATTTGATGAGAAATTAATG 400  
 DB 341 CTACTGATGATGAAACCAATTAATTAATGAGATCTTAATTAATTAATTAATTAATTA 400  
 QY 401 ATCTATGATACGATTTTACCGCTTCACTGATTAATTTGCAACGAGCGGCGCAATC 460  
 DB 401 GTGATGAATCTTCCGTTATCCCTTCATGCTGCTGCTGAT--AAAGAGCAATCTA 457  
 QY 461 TGAAGATGAACCAAGAGAGCGGCTCAAGTCTGATTAATCACTGCTGATCA 520  
 DB 468 TTGCTGTTGAAAGCAAGAAAGGCTTACGTTATTAATTAATTAATTAATTAATTAATTA 517  
 QY 521 CGAAGCGCCGCTTATTAATGATTAATCAATCTGAGCAATTAATTAATTAATTAATTA 580  
 DB 518 CAAATTAACCAATTTGATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 577  
 QY 581 CGAAGCAATTAATGAGCAAGAAAGATGAGGCGGATTAATGAGGCTTCTGCTTTGGCCAG 640  
 DB 578 CTAGGAATCTCAAAATTAATTTTTCAGATCAATTAATTAATTAATTAATTAATTAATTA 637  
 QY 641 TAGGAAGCTGTTGCTGCGGCGGATTAATTAATCAATGCTTCCGCTTTCAGAGCTGTT 700  
 DB 638 TGGGCTGATTTGGGCTTCCGAGAGATTTATCATCAATTAATTAATTAATTAATTAATTA 697  
 QY 701 ATTGGAAGCAATCTAGAGCGGCGCGATGAAAGAAAGGCTGATGACAGCGCTTTTC 760  
 DB 698 TTAATTAAGTTAATTTGATTAATCAAGAAATTAATTAATTAATTAATTAATTAATTAAT 757  
 QY 761 AAATTTGCAATTAATGAGATTAATCAAGGCGCGATTAATGAGGAGAAAGCAATTAAT 820  
 DB 758 ATATTTTAAGTTCTGTCGAACCAAAAGATTTGATGATTTGATTAATTAATTAATTAAT 817  
 QY 821 ATTAATGCAATTAATTAATGATTAATGCAAGAAATTAATTAATTAATTAATTAATTAAT 880  
 DB 818 AGATTAAGATTAATTTCTTCAATTTGATTAATGAGGAAATTTATTAATTAATTAATTAAT 877  
 QY 881 ATGACAAATCGGCAATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 929  
 DB 878 ATGACAAATGTCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 926

RESULT 8  
 US-09-107-532A-366  
 ; Sequence 366, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 NUMBER OF SEQUENCES: 7310  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 STREET: 100 Beaver Street  
 CITY: Waltham

STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02354  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: PC  
 OPERATING SYSTEM: <Unknown>  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/107,532A  
 FILING DATE: 30-Jun-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ariadello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-5007  
 TELEFAX: (781)893-8277  
 INFORMATION FOR SEQ ID NO: 366:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 978 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (1) LOCATION 1...978  
 SEQUENCE DESCRIPTION: SEQ ID NO: 366:  
 US-09-107-532A-366  
 Query Match 8.3%; Score 82.6; DB 3; Length 978;  
 Best Local Similarity 45.8%; Pred. No. 1.8e-15;  
 Matches 363; Conservative 0; Mismatches 424; Indels 6; Gaps 2;

QY 197 TAATTTGCGAGCGCAATTAATGAAAGCGTTATCTTTCGCGGCTTATTTCCGGGCT 256  
 DB 191 TTTATTAAGATGCGCAATTAATGAAAGATTAATGATGATGCTGGCTTAATTTTCTGGGT 250  
 QY 257 ATGCGAGTACCAAAAGCAATTAATGAAAGATTAATGATGATGCTGGCTTATTTCCGAGTTG 316  
 DB 251 ATGCTGATTAATTAAGAA--ATACAAAGAGGAAAGCAATGATATCTCTTTGAATTTA 307  
 QY 317 TCAGATGGGCTGCTGAGTCTGCTGCTTTGGAAGCGTAAAGAAAGATTCGATCTT 376  
 DB 308 TTCTTTGATTTTAAAGCAATGCTCAACGATGAGGAAAGCTAAATTAATTAATTAATTAAT 367  
 QY 377 TAACATTTGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 436  
 DB 368 TCAATTTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 427  
 QY 437 TGTGAGACCGGCGGCGGCACTGAGATGAACCAAGAGAGCGGCTCAAGTCT 496  
 DB 428 TAGCGAT--AAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 484  
 QY 497 ATGATTAATCACTGCTGCTGATGAGCAAGCGGCGGCTTAAATGAGCAATTAATTAATTAAT 556  
 DB 485 ATGATTAATCACTGCTGCTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 544  
 QY 557 TCGAGCAATTAATCAAGATTAATCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 616  
 DB 545 TAAACAAATTAATGATGCTTATGAGTGAATCTCTAAATTAATTAATTAATTAATTAATTAAT 604  
 QY 617 CCGTTTCTGCTTTGGCGAAGGCTTGAATGATGCTGCTGCGGCGGATTAATCAAGCGC 676

Db 605 GTTGAATGCTATAGCCGCGTATGGAGGATAGGCTTGGCTGGAATTTATCTCAG 664  
Qy 677 CTTCGCCGTTGTGACAGCTGTTTATTTGAAGAATCTAGAGCCGGCGCATGAA 736  
Db 665 TATCTCGTTTGTAAACGACTTTACGAGCTGATTCGTATCTGAGATTCAGAT 724  
Qy 737 CGAAAGGTGTAACGCCGCTTTTCAATTTTGGCAATATGACATACCAAGGCGCAG 796  
Db 725 CAGAAAGTATAGCAATTTTCCATCTTAGGTTAGTGAACACAAAGGTTGT 784  
Qy 797 TGAATACGAGAGACGAATATCATTTAGGCATATATCTCGTATGACAGAA 856  
Db 785 GTGATGTTGTGAAGAAATATGAAATATGCAATTTATCTTGTGCAATGTTGCA 844  
Qy 857 CTGAAATCTATTTCCACCATATGACATCGCAATCCAAATTAATTTATTC 916  
Db 845 AAGGATATATTTATTTGAACATATGAAACAGTCAATTTACTGCAATTTATGAA 904  
Qy 917 ATGAAGCTTGAAGCTTTGAGCGCTAAAGTGTTCGCTAAAGCAGAGAGTATTC 976  
Db 905 AAGAGCTTGAATAGTATAGTAAATTTATTTATTCATATATGAAAAACAAATTA 964  
Qy 977 ATGAGCTTAATTA 989  
Db 965 AATATATTAATTA 977

## RESULT 9

US-08-956-171E-3720/C

Sequence 3720, Application US/08956171E  
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956.171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 3720:

SEQUENCE CHARACTERISTICS:

LENGTH: 400 base pairs

TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 3720:  
US-08-956-171E-3720

Query Match 4.8%; Score 47.8; DB 3; Length 400;  
Best Local Similarity 48.1%; Pred. No. 0.0001;  
Matches 136; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

Qy 286 GATACCGTCAATTCGCCGATGAGTTGTGACATGGGCGTGTAGTGTGACGCT 345  
Db 333 GAGCGATGATATATTCGCAAAATGATGATGATTTGGATATCAACAAGC 274  
Qy 346 TTGAGACGCTTAAAGAAAGATTCATCTTAAAGATTTAGAGAAATTAATGATCTA 405  
Db 273 ATGGAATATGAAACAAAGACATCCCAATCATGTTAGCTGATATTTAATGAC 214  
Qy 406 TTGATACAGTTTACCGCTTCACTGATATTTGTGACACCGGACCGGCGGAACTGACG 465  
Db 213 ATCGTGATGTTCCGCTATTCATATCATGTTTCCGATGCACTGACATACAGTCGAA 154  
Qy 466 ATGAACCAAGAGACGCGCTCAAGTCTATGATATCACTGCTGATGACGAC 525  
Db 153 GTTCAATTTAAAGGAGTGAAGTGTATTAAGATATCTTATGATGCTTTAAACAAT 94  
Qy 526 AGCCCGACTTTATATGACATGTAACCATCTGACGACATATA 568  
Db 93 CATCCAGACTTAATTTGGCATTTATGATTTAAGCAATATA 51

## RESULT 10

US-08-781-986A-3720/C

Sequence 3720, Application US/08781986A

Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 3720:

SEQUENCE CHARACTERISTICS:

LENGTH: 400 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-3720

Query Match 4.8%; Score 47.8; DB 3; Length 400;



QY 862 AACTACTATT 872  
DB 3714 AACTATTTT 3724

## RESULT 13

US-09-453-702B-62/c

Sequence 62, Application US/09453702B  
Patent No. 6365723

## GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.  
Burland, Valerie  
Perna, Nicole T.  
Plunkett, Guy  
Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady  
STREET: 1 South Pinckney Street  
CITY: Madison  
STATE: WI

COUNTRY: US  
ZIP: 53701-2113

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B  
FILING DATE: 03-DEC-1999  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998  
ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166

## INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 61663  
TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-09-453-702B-62

Query Match 3.7%; Score 36.4; DB 3; Length 61663;  
Best Local Similarity 49.5%; Pred. No. 5.7;

Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 615 AGCCCTTCTGCTTTGGCGAAGCTTAGAAGCTGTCGCGGGGATTAACACC 674  
DB 1238 AGCCCTTCTGCTTAAGCTTAGAAGCTTGGAGATGCGACACCTTAATCGG 12879  
QY 675 GCCTTCCGGTTTGTACAGCTGTTAATTTGAAAGAACATCTAGAGCGGCGCGATGA 734  
DB 12878 CTATCAAGATTTGCAATAGAAATATCTTTTGTAAGTCTTCACTAGATTAAGTTGA 12819  
QY 735 AAGCAAGGTGTACAGCGCTTTTCAATTTTGGCAAAATATGACGATCAACAAAGGCGC 794  
DB 12818 ACTGAAAGTTTGGCAAGTGTATGATGTAAGTCTAGTTCAGAAATGAAAAGTCCGC 12759  
QY 795 AGTGATAACG 804  
DB 12758 AAGCAATTCG 12749

## RESULT 14

US-10-114-170-62/c

Sequence 62, Application US/10114170

Patent No. 6855814

## GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.  
Burland, Valerie  
Perna, Nicole T.  
Plunkett, Guy  
Welch, Rod

TITLE OF INVENTION: No. 6855814el Sequences of E. coli O157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady  
STREET: 1 South Pinckney Street  
CITY: Madison  
STATE: WI

COUNTRY: US  
ZIP: 53701-2113

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/114,170  
FILING DATE: 01-APR-2002  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/453,702  
FILING DATE: 03-DEC-1999  
APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998

## ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166

## INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 61663  
TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-10-114-170-62

Query Match 3.7%; Score 36.4; DB 3; Length 61663;  
Best Local Similarity 49.5%; Pred. No. 5.7;

Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 615 AGCCCTTCTGCTTTGGCGAAGCTTAGAAGCTGTCGCGGGGATTAACACC 674  
DB 1238 AGCCCTTCTGCTTAAGCTTAGAAGCTTGGAGATGCGACACCTTAATCGG 12879  
QY 675 GCCTTCCGGTTTGTACAGCTGTTAATTTGAAAGAACATCTAGAGCGGCGCGATGA 734  
DB 12878 CTATCAAGATTTGCAATAGAAATATCTTTTGTAAGTCTTCACTAGATTAAGTTGA 12819  
QY 735 AAGCAAGGTGTACAGCGCTTTTCAATTTTGGCAAAATATGACGATCAACAAAGGCGC 794  
DB 12818 ACTGAAAGTTTGGCAAGTGTATGATGTAAGTCTAGTTCAGAAATGAAAAGTCCGC 12759  
QY 795 AGTGATAACG 804  
DB 12758 AAGCAATTCG 12749

RESULT 15  
US-09-455

US-09-455-486-7/c  
; Sequence 7, Application US/09455486

; Sequence 7, Application US/09455486  
; Patent No. 6833438

Patent No. 6833438  
GENERAL INFORMATION:

APPLICANT: Daniel

APPLICANT: Daniel E. Afar  
APPLICANT: Rene S. Hubert

APPLICANT: Rene S. Hubert

APPLICANT: Arthur B. Raitano

APPLICANT: Douglas C. Safran

APPLICANT: Douglass C. Sullivan  
;  
APPLICANT: Stephen C. Mitchell  
;

APPLICANT: Stephen C. Mitchell  
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS

1 TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS  
2  
3  
4 TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833

FILE REFERENCE: 129.17-US-11

FILE REFERENCE: 129.17-US-11  
CURRENT APPLICATION NUMBER: IIS/09/455 486

CURRENT APPLICATION NUMBER: US/0  
CURRENT FILING DATE: 1999-12-06

CURRENT FILING DATE: 1999-12-06

COCKRENT FILING DATE: 1999-12-  
PRIOR APPLICATION NUMBER: 09/

;  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE: 1999-

```

; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 34

```

; NUMBER OF SEC  
; SOFTWARE: Fa

SOFTWARE: Fa  
; SEQ ID NO 7

SEQ ID NO	LENGTH
7	4

LENGTH: 4

TYPE: DNA

LIFE: DNA  
ORGANISM: HOMO

```

; ORGANISM: Homo sapiens
; FEATURE:

```

```

; FEATURE:
; NAME/KEY: CDS

```

LOCATION: (85)...(1464)  
US-09-455-486-7

Query Match 3.5%; Score 34.4; DB 3; Length 4429;

Query Match	3.5%;	Score 34.4;	DB 3;	Length 4429;
Best Local Similarity	51.3%;	Pred. No. 6.2;		

243 TTTATTTCCGGCTATCCGAGTACGAAAAACATACGGGAGATACCGTTACATTGT 302

Db 1726 TCAATCTAGCTCCATGGCCTCTGGAAATAAGAGTCAGGACCTGCACCTGATTCCTCA 166

303 CCGGCATGACTTTGTGACATGGGTCTGTCACTCTGTCACTTTTGGAGACGTAAAGA 362

[illegible][illegible]

27 363 AAAGATTCGATCTTAACGATGGAGGAAAAAT 398

Search completed: March 18, 2006, 11:58:33  
Job time : 228 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:52:29 ; Search time 992 Seconds

(without alignments)  
8302.722 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996  
Sequence: 1 catatgacacacagctctac.....atgagcttaattgaagatcc 996

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA\_Main.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	996	100.0	996	US-10-812-387-1	Sequence 1, Appli
2	439.8	44.2	990	US-09-974-300-1719	Sequence 1719, Ap
3	287.8	28.9	984	US-10-282-122A-9120	Sequence 9120, Ap
4	237	23.8	969	US-10-282-122A-9817	Sequence 9817, Ap
5	232.6	23.4	980	US-10-398-221-766	Sequence 766, App
6	232.6	23.4	990	US-10-398-221-2682	Sequence 2682, Ap
7	232.6	23.4	990	US-10-282-122A-24755	Sequence 24755, A
8	173.8	17.4	1616	US-10-429-802-17	Sequence 17, Appli
9	173.8	17.4	1616	US-10-430-503-8	Sequence 8, Appli
10	160.2	16.0	978	US-10-282-122A-34300	Sequence 34300, A
11	159	16.0	996	US-10-724-972A-1455	Sequence 1455, Ap
12	146	14.7	993	US-10-282-122A-7960	Sequence 7960, Ap
13	146	14.7	11466	US-08-781-986A-444	Sequence 444, App
14	146	14.7	11466	US-10-329-624-444	Sequence 444, App
15	142.4	14.3	1002	US-09-815-242-8383	Sequence 8383, Ap
16	142.4	14.3	981	US-09-815-242-4665	Sequence 4665, Ap
17	117.8	11.8	972	US-10-282-122A-35759	Sequence 35759, A
18	95.4	9.6	972	US-10-282-122A-20906	Sequence 20906, A
19	80.6	8.1	972	US-10-282-122A-21716	Sequence 21716, A
20	80.2	8.1	12256646	US-10-470-565-1	Sequence 1, Appli
21	72.4	7.3	1050	US-10-282-122A-12199	Sequence 1199, A
22	63.6	6.4	978	US-10-398-221-911	Sequence 911, App
23	63.6	6.4	978	US-10-398-221-2826	Sequence 2826, App

24	61.8	6.2	579	7	US-10-282-122A-21539	Sequence 21539, A
25	55.2	5.5	1092	7	US-10-282-122A-40070	Sequence 40070, A
26	53.6	5.4	1092	7	US-10-282-122A-38691	Sequence 38691, A
27	52.2	5.2	9749	3	US-09-070-927A-154	Sequence 154, App
28	47.8	4.8	400	2	US-08-781-986A-3720	Sequence 3720, Ap
29	47.8	4.8	400	7	US-10-329-624-3720	Sequence 3720, Ap
30	40.8	4.1	37973	6	US-10-311-455-2169	Sequence 2169, Ap
31	37.2	3.7	1068	3	US-09-815-242-6788	Sequence 6788, Ap
32	37.2	3.7	10528	6	US-10-311-455-307	Sequence 307, App
33	36.4	3.7	61663	6	US-10-418-837-1	Sequence 1, Appli
34	36.4	3.7	61663	5	US-10-118-170-62	Sequence 62, Appli
35	36	3.6	826	3	US-09-070-927A-799	Sequence 799, App
36	35.8	3.6	583	4	US-09-925-065A-392869	Sequence 392869, A
37	35.8	3.6	3673778	6	US-10-312-841-2	Sequence 2, Appli
38	35.4	3.6	245	3	US-09-880-107-1689	Sequence 1689, App
39	35.4	3.6	419	4	US-09-925-065A-654182	Sequence 654182, A
40	35.4	3.6	419	4	US-09-925-065A-654183	Sequence 654183, A
41	35	3.5	600	9	US-10-972-079-44650	Sequence 44650, A
42	35	3.5	661	7	US-10-648-553-102	Sequence 102, App
43	35	3.5	1053	9	US-10-450-763-28412	Sequence 28412, A
44	35	3.5	11036	5	US-10-239-676-118	Sequence 118, App
45	35	3.5	11036	6	US-10-240-453-138	Sequence 138, App

## ALIGNMENTS

RESULT 1	US-10-812-387-1	US-10-812-387-1
Sequence 1, Application US/10812387		
Publication No. US20050142652A1		
GENERAL INFORMATION:		
APPLICANT: SIVARAMAN, HEPHIZIBAH		
APPLICANT: PUNDLE, ARCHANA VISHNU		
APPLICANT: SURESH, CHEERAVAKATTU GOPALAN		
APPLICANT: DODSON, GEORGE GUY		
APPLICANT: BRANNINGAN, JAMES ANTHONY		
TITLE OF INVENTION: PROCESS FOR PRODUCTION OF LARGE AMOUNT OF PENICILLIN V		
TITLE OF INVENTION: ACYLASE		
FILE REFERENCE: 056859-0196		
CURRENT APPLICATION NUMBER: US/10/812.387		
CURRENT FILING DATE: 2004-03-30		
PRIOR APPLICATION NUMBER: PCT/IB03/06198		
PRIOR FILING DATE: 2003-12-24		
NUMBER OF SEQ ID NOS: 1		
SOFTWARE: PatentIn Ver. 3.2		
SEQ ID NO 1		
LENGTH: 996		
TYPE: DNA		
ORGANISM: Bacillus subtilis		
US-10-812-387-1		
Query Match	100.0%;	Score 996; DB 9; Length 996;
Best Local Similarity	100.0%;	Pred. No. 1.6e-289;
Matches 996; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
1	CATATGACACAGTCTTACATTTGAAAGCTGACCGTAAACATGATTAGCAAGACA	60
1	CATATGACACAGTCTTACATTTGAAAGCTGACCGTAAACATGATTAGCAAGACA	60
61	ATGAGATTTGCAATTCAGTGGGACAGAGGTATCTATCCGCGCTTACAGCTGG	120
61	ATGAGATTTGCAATTCAGTGGGACAGAGGTATCTATCCGCGCTTACAGCTGG	120
121	AACAGTGAAGCTACGGAAGGCGCCATCAACACAGTACCGCTTATCGGTATGGGAGACA	180
121	AACAGTGAAGCTACGGAAGGCGCCATCAACACAGTACCGCTTATCGGTATGGGAGACA	180
181	AAACTGGAATATATTTATTTGCGGACGATTAATGAAGCGTTATCTTGCGGCG	240
181	AAACTGGAATATATTTATTTGCGGACGATTAATGAAGCGTTATCTTGCGGCG	240
241	CTTATTTTCCGGGCTATGCGAGTACGAAAAAGATACGGAGATACCGTTACATT	300

```

Db      241 CTTTATTTTCCGGGCTATCGAGTACGAAAAACATACGGGAATACCGTTACATT 300
Qy      301 GTCCGCATGAGTTTGTGACATGGGTCTGTCACTGTCTGTCAGTCTTTGGAAGCTTAAA 360
Db      301 GTCCGCATGAGTTTGTGACATGGGTCTGTCACTGTCTGTCAGTCTTTGGAAGCTTAAA 360
Qy      361 GAAAAGATCGATCTTTAAGATTGTAGAGAAAAATATGATCTATTTGATCAGTTTA 420
Db      361 GAAAAGATCGATCTTTAAGATTGTAGAGAAAAATATGATCTATTTGATCAGTTTA 420
Qy      421 CCGCTTCACTGATATTTGTACAGCCGAGCGGCGGAACTCTAGCAGTAGAACCAAGACA 480
Db      421 CCGCTTCACTGATATTTGTACAGCCGAGCGGCGGAACTCTAGCAGTAGAACCAAGACA 480
Qy      481 GACGGCTTCAAGCTATGATTAATCAACTGTGTGTATGACGAAACGCCCTTTTAA 540
Db      481 GACGGCTTCAAGCTATGATTAATCAACTGTGTGTATGACGAAACGCCCTTTTAA 540
Qy      541 TGGCATGTAAACCAATCTGACAGCAATATACAGAAATCAACCGAAGCAATTAGAGACAA 600
Db      541 TGGCATGTAAACCAATCTGACAGCAATATACAGAAATCAACCGAAGCAATTAGAGACAA 600
Qy      601 GAGATGGCGGATTAAGCCCTTCTGCTTTTGCCCAAGGCTTAGAACTGTGTGCTGCC 660
Db      601 GAGATGGCGGATTAAGCCCTTCTGCTTTTGCCCAAGGCTTAGAACTGTGTGCTGCC 660
Qy      661 GGGGATTAATACACCGCTTCCCGGTTTGTACAGCTGTATTTGAAAGAACTATAGAG 720
Db      661 GGGGATTAATACACCGCTTCCCGGTTTGTACAGCTGTATTTGAAAGAACTATAGAG 720
Qy      721 CCGGCGGCGGATTAAGCAAGGATGTACAGCCGCTTTCAAAATTTTGGCAATATAGAG 780
Db      721 CCGGCGGCGGATTAAGCAAGGATGTACAGCCGCTTTCAAAATTTTGGCAATATAGAG 780
Qy      781 ATACCAAGGCGCAGTATGATACGAGAAAGCAAGAAATTCATTATACGCAATATCTCC 840
Db      781 ATACCAAGGCGCAGTATGATACGAGAAAGCAAGAAATTCATTATACGCAATATCTCC 840
Qy      841 GTGATGTGCAACGAACTGGAATCTATTTTCCACACTATGACAAATGGCAATTCGA 900
Db      841 GTGATGTGCAACGAACTGGAATCTATTTTCCACACTATGACAAATGGCAATTCGA 900
Qy      901 AAGTAAATTTATTTCAAGAACTGACGTTTGGAGCTTAAAGTCTTTCCGCTAAA 960
Db      901 AAGTAAATTTATTTCAAGAACTGACGTTTGGAGCTTAAAGTCTTTCCGCTAAA 960
Qy      961 GCAGAGAGATTCATGAGCTTAATTAAGATCC 996
Db      961 GCAGAGAGATTCATGAGCTTAATTAAGATCC 996

RESULT 2
US-09-974-300-1719
; Sequence 1719, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 1085, 500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1719
; LENGTH: 990
; TYPE: DNA

```

```

; ORGANISM: Bacillus licheniformis
US-09-974-300-1719
Query Match      44.2%   Score 439.8; DB 3; Length 990;
Best Local Similarity 65.3%; Pred. No. 1.5e-121;
Matches 645; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

Qy      4 ATGTCACAGCTTTACATTTGAAACTGCTGACCGTAAACATGATATAGCAAGAACATG 63
Db      1 ATGTCACAGATATCACTGTGACGACTCGGGCCGGAGCATTTATTTGGCGAGACATG 60
Qy      64 GATTTGCAATTTCACTGGGAGCAGAGTGAATCTCTATCCGCCGCTTACAGCTGAAAC 123-
Db      61 GATTTGCAATTTGACCTTAAACGGGAGGTGTGCTGTGACATCCGCTGTACAAATGAAA 120
Qy      124 AGTGAAGTGAACGGAAGGCCCATACAGACAGTACGGTTATATGATATGAGGAGAAA 183
Db      121 AGTGAAGCGGAGCGGAGCGAGCATCCCGCGGTATGCGTTTATCGGATGGGAGAAAG 180
Qy      184 CTGGAATAATATATTTATTTGCGCAGCGCATTAATGAAAGCGTTTATCTTGTCCGCGCTT 243
Db      181 CTCAGAAATGCTCTGTTTGCAGATGTGTGATGAAAGATGATGCTTGCAGCGGCGCTT 240
Qy      244 TATTTTCCGGGCTATGCGGAGTACGAAAAACGATACGGGAATACCGTTGCAATGTC 303
Db      241 TATTTTCCGGGCTATGCTGTGTATGAAAAAGCAAGCAAGCAAGCCGCAATCTTGC 300
Qy      304 CCGCATGAGTTGTGACATGGGTGCTGTGATCTGTGATCTTTGGAAAGCTTAAAGAA 363
Db      301 CCGCAAGATTTGTGACATGGGTGCTGTGATCTGTGATCTTTGGAAAGCTTAAAGAA 360
Qy      364 AAGATTCATCTTTAAGATGTGAGAAAAATTAATGATCTATGATACAGTTTACCG 423
Db      361 GGGGCTTTCTTAAACATTTGTGAGAGAGATCACTGCTGATGATCGGTGACCGCG 420
Qy      424 CTTCATGATATTTGTACACCGGAGCGGCGGAAACCTGACATGAAACCAAGAGAGAC 483
Db      421 CTGCATGTCTATTTAAGCAAGCGCTCCGGGCAAGGTGTGTTGATCCGAGGAGAC 480
Qy      484 GGGCTCAAGTCTATGATATCAACTGTGTGTATGACAGCAAGCCCGCATTTATATAG 543
Db      481 GGGATTCAAATTCATATATATCCGGTCCGCTGTCTACCAAGCTCCGATTTCCCTTGG 540
Qy      544 CATGTAAACCAATCTGACGAAATATACAGAAATCAACGGAAGCAATTAAGAGCAAGAG 603
Db      541 CATTTACCAATTTGGCAACTTATCCGCTTCAACCGGAGCAATTTGCCGCAAAAAA 600
Qy      604 ATGGGCGATTAAGCCCTTCTGCTTTTGGCCAAAGCTTATAGAACTGTGTCTGCCGGGG 663
Db      601 ATGGGAGGCTGACGCTGTCTGCTGTGTGCAAGGTTCCGGGCTGTCCGATTTGCAAG 660
Qy      664 GATTATACACCGCTTCCCGGTTTGTCAAGCTGTATTTTGAAGAAACATCTAGAGCCG 723
Db      661 GATTTTACCGCGCATTCGCTTTGTCAAGGCTGTATTTTGAAGAAACATCTAGAGCCG 720
Qy      724 GCGGCGGATGAAACGAAGGTGTACAGCGCTTTTCAAAATTTTGGCAATATAGAGATA 783
Db      721 GTGTCCGATGAAACGAAGGCTGTCTGCGGCTTTTCAATGTCTCTCCATATAGAAATTT 780
Qy      784 CCAAGGCGCAGTATGATAGGAAAGCAAGAAATTCATATACGCAATATACCTCGGT 843
Db      781 CTTAAGGGGTGTGTATGATGATCAAGGGGAAGCAATTAACCCCAATATACGGCGTA 840
Qy      844 ATGTCAACGAACCTGGAATCTACTATTTTCCACCATATGACATCGGCAATTCAAAAA 903
Db      841 ATGTCAACGAACCTGCACTACTATTTTCCACCATATTAACAAACGCGAGATTCAAAAG 900
Qy      904 GTAATTTATTTATGAAAGCTTGAAGCTTGTGAGCTTAAAGTGTCTTCCGCTAAAGCA 963
Db      901 GTTTGCTATTTACAGAAAGCTGACAGCAAGAAATTAAGGCTTTTCAAGCAAGACAG 960
Qy      964 GAAGAGATTCATGAGCTTAATTA 990

```



DB 961 GAAGAGCGTTCTTCTTCATCA 987

RESULT 3

US-10-282-122A-9120

Sequence 9120, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Foreyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 9120

LENGTH: 984

TYPE: DNA

ORGANISM: Bacillus anthracis

US-10-282-122A-9120

Query Match 28.9%; Score 287.8; DB 7; Length 984;

Best Local Similarity 56.0%; Pred. No. 1.2e-75;

Matches 544; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

DB 4 ATGTGCAAGTCTTACATGGAAGCTGCGTAACGATGATTAAGCAAGAACATG 63

1 ATGTGTAATGCTTGTGACATTAAGCAAAAAACGCTGACATCTTTTGGCAAGAACATG 60

64 GATTTGCAATTCAGCTGGGACACAGAGTGAATTCCTATCCGCGCCCTTAAGCTGAGAC 123

61 GACTTCACATTAAGTATGATGATCAAGAGTATATATCTTCTGACATTTCCAGTGAAT 120

124 AGTGAAGCTACGGAAGGCGCCATCAACACAGTACCGCTTTATCGTTATGGGAGAAA 183

121 AATATTAACGGGTGAATCAATTAACCAAAACATGCTACGGTCGAATGGGTATTAATCAT 180

184 CTGGAATAATATATTTGCGACGAGCATTAATGAAGCGTTATATCTTGGCGCGCT 243

181 CAAGGAAGATCATATATGGCGACGAGTAAATGAACAGGTATGACATGTGCAACATC 240

DB 244 TATTTCCGGGCTATGGGAGTACAAAAACGATACGGGAAGTACCGTTCACTTGTG 303

241 TATTTCCAGGATTCGCTACTTATAGTCAAAAGCATAGTGCACACAGCAATTTGCT 300

304 CCGCATGATTTTGTGACATGGGTCCTGTCACTGTCTGATCTTTGGAAAGCGTAAAGAA 363

301 CCATTTGATTTTGTGATCTTGGAGTGTGACCAATTCATCTGTCAAAAGTAAAGAA 360

364 AAGATGATCTTAAAGATTTGTAGAGAAAAAATTAATCATATGATGATACGTTTACCG 423

361 TCTGTATATGATACCTTCTTTTGGATATACATTAACCGATTTAGACTTAACGCCACA 420

424 CTTCATGATATTTGTGACACCGGACGGCGCAAACTTGACATGATGAACCAAGACAGAC 483

421 CTACATTTGATTTTGGAGGATTAATGGGAGATTTGACATTTGATGATGACAGATGA 480

484 GGCTCAAGTCTATGATTAATCAACTGTGTGTGACGAAACAGCCCGACTTATATG 543

481 GGATTAATAATTTGATGATTAACCCACTGAGGTGATGACGAATATGTCCGAGTTTAAATGG 540

544 CATGTACCAATCTGACGAATATGAGGATGACGAAAGCAATTAAGAGCAAGAG 603

541 CATTTACAAATTTAAGACAAATATATAGGCTTTAATGCAAGCCATTCGGCAACAGAG 600

604 ATGGCGGATTAAGCCCTTTCTGCTTTTGGCCAAAGCTTGAAGAACTGTGTCTGCGGG 663

601 TGAGGATATTTTACATTAATGCTTTTGGCCAAAGCTTGAAGAACTGTGTCTGCGGG 660

664 GATTAATACCGCTTCTCCGCTTTGTGACAGCTGTGTTATTTGAAGAACATCTAGACCG 723

661 GATTTCAACCCCGCATGAGGTGTTGTGCGGACGACATATGCAAAACAAATCAAGT 720

724 GCGGCGGATTAAGCAAGGTGTAACGCGCTTTCAATTTTGGAAATATATACATA 783

721 ATGATATGCAAGAGGAGGATATGACGCTTTTCAATTTTGAATATGAGGTT 780

784 CCAAGGCGCAGTATTAACGAAGAAAGCAAAATTCATTAACGAATATATCTTCCGTG 843

781 CTAAGGTGATGATTAATTAAGAGAGAGTGCATTAATTAATTAATTAATTAAGGTA 840

844 ATGTGCAAGAACTGGAACCTATTTTCAACCATATGACATGACCAATTCGCAATCAAAA 903

841 ATGTGATGGAATCCGGAACATATTTATCATATCTTACATTTGATGACAAATATAGCT 900

904 GTAATTTATTTATGAGACCTTGAATGCTGTTTGAAGCTTAAGCTTTCCGTAAGCA 963

901 GTTCATTTATTTATGAAATTTAATTAAGATGATTAAGCTTATCCGTTCAACGG 960

964 GAAGAGATAT 974

961 AACAATAAT 971

RESULT 4

US-10-282-122A-9817

Sequence 9817, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Foreyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

Query Match	23.8%	Score 237;	DB 7;	Length 969;
Best Local Similarity	55.5%	Pred: No. 2.5e-60;		
Matches 538; Conservative	0;	Mismatches 425;	Indels 6;	Gaps 4

Oy 4 ATGAGCAAGAGCTTACATTTGGAAATGCTGACCGTAAACATGATTTACCAAGCAATG 63  
 Db 1 ATGTGTACTAGTTTGTGCATTTAGAGCAAAAAAGGTACAGATCTTTTTCAGAACGATG 60  
 Oy 64 GATTTTGCATTTTCAGCTGGGGACAGAGTGATTTCTTATCCGGCCGTTTACAGCTGGAAC 123  
 Db 61 GACTTCACATTTGATATGATATCAAGATTAATATCATTTCTGTGACATTACAGTGGAAAT 120  
 Oy 124 AGTGAAGCTGACGGAAAGGGCCCATCAGACACACTACCGCTTATCGGTATGGGAGAAAA 183  
 Db 121 AATATTAACGGGTGAATCATTTAATACGAAACATGCTACGGTCCGAATGGGTATTAATCAT 180  
 Oy 184 CTTGGAATATATATTAATTTGCCGACGGCATTAAGAAAGGGTTTATCTGTGCGGGGCTT 243  
 Db 181 CAGGAAGGATCATTTATGCGGACGGATTAAGAAACAGTATGACATGTGCACATCTC 240  
 Oy 244 TATTTTCCGGGCTATCCGAGTACGAAAAAAAGATACGGGAAGATACCGTTTCATTTGCT 303  
 Db 241 TATTTTCCAGATTCGCTACTTATTACTCAAGCATATGATACACACACACGATTTGGCT 300  
 Oy 304 CCGCATGAGTTTGTGCATGAGGTGCTGTCACTGTCACTTTT--GGAAGACGTAAAG 361  
 Db 301 CCATTTGATTTTGTACTTGGAGCTGACACCAATTCATTTCTGTCAAGGCCAGTGAAG 360  
 Oy 362 AAAAAGTTGATTTTAAACGATTTGTAGAAAAAATTTATGATCTAATGATATAC--AGTTTAA 420  
 Db 361 AATCTGTAGATGACATTAACCTTTTGGATATACCATTTACCGGCAATTTAGACATTAGCCCA 420  
 Oy 421 CCGCTTCACGTGATTTTGTGCAGACCGGACGGGCGGAAACCTGTACGATAGAACCAAGACA 480  
 Db 421 CCACATCATTTGATTTTAAAGCGATTAATATGGGAGATTTGATGTATCGATCCGACAGT 480  
 Oy 481 GACGCGCTCAAGCTTATGATTAATCAACTGTGTCTATGACGAACAGCCCGGACTTTATA 540  
 Db 481 GAAGGATTAATAATTTGATATGATTAACCCACTGGAGTGATGACGAATAGTCTCGGAGTTTAA 540  
 Oy 541 TGGCATGTAAACCATTCGACAGCAATATACAGAAATGACACGAAAGCAATTAAGAGACAAA 600  
 Db 541 TGGCATTTACAAATTTTAAGACAAATATATAGGCTTAAATGCGACGCATTTGGCCACA 600

QY	601	AAAGATGGCGCGATTAGCCCTTCTGCTTTTGCCCAAGGCTTAGGAACGTGTGCTGCCG	660
Db	601	GAGTGGATAATTATTCACATTAAGTCTTTTGCCCAAGGCTCGGCTCAATGGACCTTCCA	660
QY	661	GGGATTTATACACGCGCTTCCCGTTTGTCAGAGCTGTATTATTGAAGAACAATCTAGAG	720
Db	661	GGGATTTTCAACCCCGCCATCCAGTGTGTGGGGGAGCATATGGCAAAACAAACATTCAA	720
QY	721	CCGGCGGCCGATGAACGAAAGGTGTACAGCCGCTTTTCAATTTTGGCAAAATATAGAG	780
Db	721	GGTATAGTATACCGAAGAGGGAGTATCAGCCCTTTTTCATATCTTATCAAAATGTGAG	780
QY	781	ATACCAAGGCGCGATGATTAACGGAAGAAAGCAAAATTCATTATATACGCAATATA-CTTC	839
Db	781	GTTCTTAAGGGAGTATATACAGAACAGGTGCATTAGTATATATATATATATACCAAG	840
QY	840	CGTATGTGCACAGAACTGGAACCTACTATTTCACCACTATGACAAATCGG--CAATC	897
Db	841	CGTATGTGATGAGAAATCCGGAACTATTTATCTTATCTTATCAATGTAGCACAATTA	900
QY	898	CAAAAAGTAAATTTATTTCATGAGAAAGCTGACGTGTGGAAACCTAAAGTGTTCGGCT	957
Db	901	TAGCTTGTCTATTATTATTTCTGAAAATTATAGTACCAGATGAATATMAACCTTATCCCTTC	960
QY	958	AAAGCAGAA 966	
Db	961	CCACCCGAA 969	

```

1      RESULT 5
2      US-10-398-221-766
3      Sequence 766, Application US/10398221
4      Publication No. US20040018514A1
5      GENERAL INFORMATION:
6      APPLICANT: KUNST, Frederik
7      APPLICANT: GLASER, Philippe
8      TITLE OF INVENTION: Listeria innocua, gen
9      FILE REFERENCE: 344 702 - US
10     CURRENT APPLICATION NUMBER: US-10/398, 221
11     CURRENT FILING DATE: 2003-03-27
12     PRIOR APPLICATION NUMBER: PCT/FR 01/03 06
13     PRIOR FILING DATE: 2001-10-04
14     PRIOR APPLICATION NUMBER: FR 00/12 697
15     PRIOR FILING DATE: 2000-10-04
16     NUMBER OF SEQ ID NOS: 4025
17     SOFTWARE: PatentIn version 3.0
18     SEQ ID NO 766
19     LENGTH: 990
20     TYPE: DNA
21     ORGANISM: Listeria monocytogenes-EGD
22     US-10-398-221-766

```

Query Match	23.4%	Score 232.6;	DB 7;	Length 990;
Best Local Similarity	53.0%	Pred. NO. 5.5e-59;		
Matches 496;	Conservative	0;	Mismatches 439;	Indels 0;
				Gaps 0;

QY 4 AATGACACAAAGCTTACATTGGAAACATGCTGACCGGAAACATGATATAGGACGAACATG 63  
Db 1 AATGACACAAAGCTTCCGATATTAGAAAGCTTAGAGGAAACATTTATTTATCAGAAACGATG 60  
QY 64 GATTTCGATTTTCAGCTGGGGACAGAGTGATCTCTATCCGGCCGCTTACAGCTGGAAAC 123  
Db 61 GATTTTGCCTTTATTTTGGAGGCGAACCACAACATTAGCCCAAGAAATTATGATGGAAA 120  
QY 124 AGTGAAGCTGACGGAAGGCGCCATCAGACACAGTACGCGTTATCGGTATAGGGAGGAAAA 187  
Db 121 TCTTCTACAGACGCTGCCAATTATATTAATAGCATTTGTGGAGCGGGAAGAGAG 184  
QY 184 CTTGGAATAATATATTTGGCCGACGCGCATTAATGAAGCGGTATCTTGGCGCGCCTT 243  
Db 181 TTAGATTAATACATTTTTCGACGCGGATTCGAATGAAAGAGATTAGCTGTGATCTCTTA 244

QY 244 TATTTCCGGCTATGCGGAGTACGAAAAACGATACGGAAAGATACCGTTCACTTCTC 303  
 DB 241 TATCTTCAGAGAGAGAGATTTATGCGCAGACACGAGGAAAGAAAAATCAATTACCT 300  
 QY 304 CCGCATGAGTTTGTGACATGGGTGCTGCACTCTGCACTCTTTGGAAAGSTAAAGAA 363  
 DB 301 CCACAGAAATTTTCTGCTTTGGCTGTAGAACTTCCGAAAGATTTAAAGATGTGAGCA 360  
 QY 364 AAGATTCGATCTTAAACGATTTAGAGAAAAATTAATGATCTATGATACAGTTTACCG 423  
 DB 361 AAGTATCGTATTAATCTTGATGATCAACAGATACCGTTACTTGAATCAACAACCA 420  
 QY 424 CTTCAGGATATTTGTGACCGGACGGCCGAAACCTGAGATTAAGACCAAGCAGAC 483  
 DB 421 TTACATGATTTTTCACAGACAGAGTGGCGTGTGTATGATTTGAACCAACAGAAACA 480  
 QY 484 GGCCTCAAGCTATGATATCAACCTGTGTCAATGACGAAACAGCCCGCATTTATATG 543  
 DB 481 TCGCTTCGTATTAAGAAACCTGTGAGATTTATGAAATACGCCCATGCAATGG 540  
 QY 544 CATGTAACCAATCTGACAGCAATATACAGAAATCAGACCGAACAATTAGAGCAAGAG 603  
 DB 541 CATATGAGATTTTACGCACTACACAGTTTACAGCAACAAATTAGCCCGGTAGAG 600  
 QY 604 ATGGGCGATTAAGCCCTTTCTGCTTTTGGCCAGGCTTGAAGCTGTGCTGCCGGG 663  
 DB 601 TTCCGAGATATATGCTTAAACCAATTTCCGAGGTACAGCAACAGCAATTTACAGGT 660  
 QY 664 GATATACACCGCTTCCCGGTTTGTGACAGCTGTATTTTGAAGAACATCTAGACCG 723  
 DB 661 GGTATATCTCCGCAAGACGTTTGTGCTGCGCATATCTTGAAGAAATATTTATAAA 720  
 QY 724 GCGGCGCATGAAACGAAAGGTGTAACGCCCTTTCAATTTTGGCAATATATGACATA 783  
 DB 721 GCGAAAAACGAAAGAAAGATATCAAAATGTGTATGTCTTGAACAGTGTCCGATT 780  
 QY 784 CCAAGGGCGCATGATACGAAAGAAAGCAAAATTCATTATACGCAATATATCTTCGCTG 843  
 DB 781 CCGATGAGACGCTTAAATTAAGAAAGTGTGACCTGATTTTCAACAATATGTGGCATCT 840  
 QY 844 ATGTCACAGAACTGGAACCTATATTTCCACCATATGACATTCGCAAAATCCAAAA 903  
 DB 841 ATGTGTCGAGAAACAAACGTACTATTTCACTCATACGAAACAAACGAAATCAATGT 900  
 QY 904 GTAAATTTATTCATGAAGACCTTGACTGTTTGA 938  
 DB 901 GTTACTTTGTCAAGAAAGTATATGAAATATGAA 935

RESULT 6  
 US-10-398-221-2682  
 ; Sequence 2682, Application US/10398221  
 ; Publication No. US20040018514A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KINIST, Frederik  
 ; APPLICANT: GLASER, Philippe  
 ; TITLE OF INVENTION: Listeria innocua, genome and applications  
 ; FILE REFERENCE: 344 702 - US  
 ; CURRENT APPLICATION NUMBER: US/10/398,221  
 ; PRIOR FILING DATE: 2003-03-27  
 ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
 ; PRIOR FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: FR 00/12 697  
 ; NUMBER OF SEQ ID NOS: 4025  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2682  
 ; LENGTH: 990  
 ; TYPE: DNA  
 ; ORGANISM: Listeria monocytogenes EGDe  
 US-10-398-221-2682

Query Match 23.4%; Score 232.6; DB 7; Length 990;

Best Local Similarity 53.0%; Pred. No. 5.5e-59;  
 Matches 496; Conservative 0; Mismatches 439; Indels 0; Gaps 0;  
 QY 4 ATGTCACAAGCTTACATTTGGAACATGCTGACCGGTAAACATATATGACAAACATG 63  
 DB 1 ATGTGACAGAGCTTCGATTTAGAAACGTTATGATGAAACCAATTTATATCAAGAACATG 60  
 QY 64 GATTTTCATTTCACTGTTGGGACAGAGGTGATTTCTATCCGCCCTTACAGCTGAAAC 123  
 DB 61 GATTTTCCTTTATTTTGGAGGGGACCCCAACATTTAGCCCAAGAAATTTATGATGAA 120  
 QY 124 AGTCAGCTGACGGAAGGCCCATCAGACAGTACCGCTTATCGGTATGCGGAGAA 183  
 DB 121 TCTTCTACAGACGGTGCATTTATATATATGATATCATTTGTTGAGCGCGGAAAGAG 180  
 QY 184 CTTCGAAATATATATTTGCGGACCGCATTAATGAAGCGGTATATCTTGCGCGCGCTT 243  
 DB 181 TTATATTAATATCATTTTTCGAGACGATGATGATGAAAGAGATTAAGCTGTGCAATCTCTA 240  
 QY 244 TATTTCCGGCTATGCGGAGTACGAAAAACGATACGGAAAGATACCGTTCACTTCTC 303  
 DB 241 TATCTTCAGAGAGAGAGATTTATGCGCAGACACGAGGAAAGAAAAATCAATTACCT 300  
 QY 304 CCGCATGAGTTTGTGACATGGGTGCTGCACTCTGCACTCTTTGGAAAGSTAAAGAA 363  
 DB 301 CCACAGAAATTTTCTGCTTTGGCTGTAGAACTTCCGAAAGATTTAAAGATGTGAGCA 360  
 QY 364 AAGATTCGATCTTAAACGATTTAGAGAAAAATTAATGATCTATGATACAGTTTACCG 423  
 DB 361 AAGTATCGTATTAATCTTGATGATCAACAGATTTATGAAATACGCCCATGCAATGG 420  
 QY 424 CTTCAGTGAATTTGTGACACCGGACGGCCGAAACCTGACATGATGAACCAAGACAGAC 483  
 DB 421 TTACATGATTTTTCACAGACAGAGTGGCGTGTGTATGATTTGAACCAAGAAACA 480  
 QY 484 GGCCTCAAGCTATGATTAATCAACCTGTGTCAATGACGAAACGCCCGCATTTATATG 543  
 DB 481 TCGCTTCGTATTAAGAAACCTGTGAGGATTAATGCAAAATGCGCACCATCTGAATGG 540  
 QY 544 CATGTAACCAATCTGACAGCAATATACAGAAATCAGACCGAATATAGAGCAAGAG 603  
 DB 541 CATATGAGATTTTACGCACTACACAGTTTACAGCAACAAATTAGCCCGGTAGAG 600  
 QY 604 ATGGGCGATTAAGCCCTTTCTGCTTTTGGCCAGGCTTGAAGCTGTGCTGCCGGG 663  
 DB 601 TTCCGAGATATATGCTTAAACCAATTTCCGAGGTACAGCAACAGCAATTTACAGGT 660  
 QY 664 GATTAATACCGCTTCCCGGTTTGTGACAGCTGTATTTGAAGAACATCTAGACCG 723  
 DB 661 GGTATATCTCCGCAAGACGTTTGTGCTGCGCATATCTTGAAGAAATATTTATAAA 720  
 QY 724 GCGGCGCATGAAACGAAAGGTGTAACGCCCTTTCAATTTTGGCAATATATGACATA 783  
 DB 721 GCGAAAAACGAAAGAAAGATATCAAAATGTGTATGTCTTGAACAGTGTCCGATT 780  
 QY 784 CCAAGGGCGCATGATACGAAAGAAAGCAAAATTCATTATACGCAATATATCTTCGCTG 843  
 DB 781 CCGATGAGACGCTTAAATTAAGAAAGTGTGACCTGATTTTCAACAATATGTGGCATCT 840  
 QY 844 ATGTCACAGAACTGGAACCTATATTTCCACCATATGACATTCGCAAAATCCAAAA 903  
 DB 841 ATGTGTCGAGAAACAAACGTACTATTTCACTCATACGAAACAAACGAAATCAATGT 900  
 QY 904 GTAAATTTATTCATGAAGACCTTGACTGTTTGA 938  
 DB 901 GTTACTTTGTCAAGAAAGTATATGAAATATGAA 935

RESULT 7  
 US-10-282-122A-24755  
 ; Sequence 24755, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:

```
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 24755
LENGTH: 990
TYPE: DNA
ORGANISM: Listeria monocytogenes
US-10-282-122A-24755

Query Match      23.4%; Score 232.6; DB 7; Length 990;
Best Local Similarity 53.0%; Pred. No. 5.5e-59;
Matches 496; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

QY 4 ATGTGCAAGTCTTACATTTGAAAGTGTGACCGTAAACATGATTAGCAAGAACATG 63
DB 1 ATGTGCAAGTCTTACATTTGAAAGTGTGACCGTAAACATTTATTATCAAGAACATG 60
QY 64 GATTTTGCATTTGAGCTGGGAGACAGAGTATCTTATCCGCCGCTTACAGCTGAGAC 123
DB 61 GATTTTGCATTTGAGCTGGGAGACAGAGTATCTTATCCGCCGCTTATGATGAGAAA 120
QY 124 AGTGAAGTCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 183
DB 121 TCTTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 184 CTGGAAGATATATATTTGCGACGAGCATTAATGAAAGCGGTTATCTTGGCGGCGCTT 243
DB 181 TTGATATATATATATTTGCGACGAGCATTAATGAAAGCGGTTATCTTGGCGGCGCTT 240
QY 244 TATTTTCCGGGCTATGCGAGTACGAAAGGAGTACGAGGAGTACCGTTACATTTGTC 303
DB 241 TATCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 304 CCGAGTAGTTTGAATGGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
DB 301 CCACAGAGATTTTGTGCTTGGCTTTGAGACTTGGCCAGATTAAGATGTGAGAGCA 360
```

```
QY 364 AAGATTCGATCTTTAAGCATTTGTAGAGAAATTAATGATCTATTTGATACATTTTACCG 423
DB 361 AAGTATCGGTTATTAATCTTGGTGTATCAACAGATACCGTTACTTGAATCAACAACCA 420
QY 424 CTTCAGTGAATTTGTGACAGCGGAGCGGAGAAACGTGACGATGAACAGAGAGAC 483
DB 421 TTACACTGATTTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 484 GGCCTCAAGTCTATGATATCAACCTGTGTGATGACAGAGAGGAGGAGGAGGAGGAG 543
DB 481 TCGCTTGTATTAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 544 CATGTAAACCAATCTGACAGATATTAAGAGATCAAGAGGAGGAGGAGGAGGAGGAG 603
DB 541 CATATCGAGATTTTACAGCACTACAGAGGTTTACAGAGCAACAATTTAGCGCGTGAAG 600
QY 604 ATGGCGGATTAAGCCCTTTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 663
DB 601 TTGAGAGATATATGCTTAACCATTTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 664 GATATATACAGCGCTTCCGCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 723
DB 661 GGTATATCTCCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 724 GCGGCGGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 783
DB 721 GCGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 784 CCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 843
DB 781 CCGAATGAGGCGGATTAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 844 ATGTGCAAGAACTGAGAACTACTATTTTCAACCACTATGACAGATGCGAATCAAAA 903
DB 841 ATGTGCGAGAAAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 904 GTAATTTATTTGATGAGAGAGCTTGTGCTTTGGA 938
DB 901 GTTACTTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 935

RESULT 8
US-10-429-802-17
Sequence 17, Application US/10429802
Publication No. US20030228285A1
GENERAL INFORMATION:
APPLICANT: HUNG, MIEN-CHIE
APPLICANT: WONG, KA YIN
APPLICANT: ZOU, YIYU
TITLE OF INVENTION: BIPARTITE T-CELL FACTOR (TCF)-RESPONSIVE PROMOTER
FILE REFERENCE: UTSC:752US
CURRENT APPLICATION NUMBER: US/10/429,802
CURRENT FILING DATE: 2003-05-05
PRIOR APPLICATION NUMBER: 60/377,672
PRIOR FILING DATE: 2002-05-03
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 1616
TYPE: DNA
ORGANISM: Bacillus sphaericus
US-10-429-802-17

Query Match      17.4%; Score 173.8; DB 6; Length 1616;
Best Local Similarity 49.8%; Pred. No. 4.1e-41;
Matches 495; Conservative 0; Mismatches 492; Indels 6; Gaps 2;

QY 7 TGCACAGTCTTACATTTGAGAACTGACCGTAAACATGATATTAGCAAGAACATGAT 66
DB 335 TGCAGTAGCTTATCAATTTGTCACAGATGATTAAGTTATTCGCTCCACAGATGAT 394
QY 67 TTGCACTTCAAGTGGGAGCAGAGGATCTCTATCCGCGCGGTTACAGAGTGAACAGT 126
```

```

Db      395 TTACATGGAACCAAGTAGTAGTATATGTCCTCCAGTAATTACGATTCGATTG 454
Qy      127 GAAGCTGACGAA---GGGCCATCAGACACAGTACGCGTTATTCGTAATGGGAGAA-- 181
Db      455 TTAGAAAAAGAAATAGTATGATTAACAATTCAATATGCTTTGTTGTAATGGGACACAT 514
Qy      182 -AACTGGAATATATATTTGGCCGAGCGCATTAAGAAAGCGCTTATCTGTGGGG 240
Db      515 GACATTAATCAACCGCTTCTATATGATGGGATTAAGAAAGGATTAATGGCCGCAATG 574
Qy      241 CTTTATTTTCCGGGCTATCGGAGTACGAAACAAATCATGCGGAGATACCGTTCAATT 300
Db      575 CTTTACTATGCTACATTTGCGACTATATGCTGACGAACCTTAAAGGACCAACAGGATC 634
Qy      301 GTCCCGCATGAGTTGTGACATGGGTGCTGCTGATCTGCTGAGTCTTTGGAGAAGTAA 360
Db      635 AATCCGCTGATGATATTTCTGATGTTAGAAATTTGTTGTAATCTGATGATGTTAT 694
Qy      361 GAAAGATTCGATCTTTACGATTTGAGAAATTAATGATCTATTTGATACAGTTT 420
Db      635 GAAATTTTACTTTCTTATACATTTATTAAGTAAGGCGCAATATATATCTTGGCTTGCACC 754
Qy      421 CCGCTTCACTGATATTTGACAGCGGACGCGCCGAAACCTGACGATAGAACCAAGACA 480
Db      755 CCACTTCACTATACATTTTACAGATGCTTCTGATGATGATGATTTGAACCGGATAA 814
Qy      481 GAGCGCTCAAACTCTATGATATCAACTGCTGCTGATGACGAAACGCGCCGACTTTTA 540
Db      815 ACAGGCATTAACCATTCATCGAATAAAGATTTGCGTATGACGATATAGCCCTGGTATGA 874
Qy      541 TGGCATGTAACCAATCTGACGAAATATACAGAAATACAGACGAAATTAAGAGCAA 600
Db      875 TGGCATGACGAAATTTAAGCTTATGATGTTGTCACCAAAATCCGCAACAGATTA 934
Qy      601 GAGATGGCGGATTAAGCTTCTGCTTTTGGCCAGGCTTGAAGTATGTTGGCTGCG 660
Db      935 ATGATGGGAGACTTGGATTTGACACCGTTTGGGCAAGGCGGCGCTTGAAGTATCA 994
Qy      661 GGGGATTAACACGCGCTTCCGCTTGTGTCAGAGCTTTTAAAGAAACATCTAG 720
Db      995 GGTGATTTTACGCGCTGACGAGCTTTCTTGGGATGACTGAAAAAATATATCTAA 1054
Qy      721 CCGGCGCGCGGATTAAGAAAGATGTAACAGCGCTTTTCAATTTTGGCAATATGAC 780
Db      1055 AAAGCAAAATGAAACAGAGGCTTAACAATCTTTCATATCTTCTCTGTAAT 1114
Qy      781 ATACCAAGGCGCGAGTATACGAGAAAGCAAAATTCATTAACCAATATATCTTCC 840
Db      1115 ATCCCAAAAGGCTTGTGTTTGAACAATGAGGGAAGAAACGATTAACATCTATACCTCA 1174
Qy      841 GTGATGTGCAAGAACTGGAATCTATATTTCCACATATGACAAATCGCAATCCAA 900
Db      1175 GCTATGTGTGCAAAAGTAAATTAATCTTAACTGTAAGCAATATGCAATTTCA 1234
Qy      901 AAAGTAATTTATTTCAAGAGACCTTACCTTTTGGAGCTTAAAGTCTTTCCGCTAA 960
Db      1235 GCCGCTTCTTAAATGGCGAAATTTAATATGTCAAAGTTTAACTTGAAGTGGAT 1294
Qy      961 GCAGAGAGATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 993
Db      1295 CGTAACAGATTTAAGCAATTAATCAAGTA 1327

```

```

RESULT 9
US-10-430-503-8
; Sequence 8, Application US/10430503
; Publication No. US200400056841
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: LAN, KENG-LI
; APPLICANT: OU-YANG, FU
; APPLICANT: LIU, JAM-CHING
; APPLICANT: LAN, KENG-HSIN

```

```

; TITLE OF INVENTION: TARGETING PROTEINS TO DELIVER THERAPEUTIC OR DIAGNOSTIC
; FILE REFERENCE: REAGENTS
; FILE REFERENCE: UTSC:79705
; CURRENT APPLICATION NUMBER: US/10/430,503
; PRIOR APPLICATION NUMBER: 2003-05-06
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1616
; TYPE: DNA
; ORGANISM: Bacillus sphaericus
US-10-430-503-8

Query Match      17.4% Score 173.8; DB 6; Length 1616;
Best Local Similarity 49.8%; Pred. No. 4,1e-41;
Matches 495; Conservative 0; Mismatches 492; Indels 6; Gaps 2;

Qy      7 TGCACAGCTCTTACATTTGGAATGCTGACCGTAACATGTAATGCAAGAACATGAT 66
Db      335 TGCAGTACTTATCAATTCGTACAAAGATGTAATTAATGCTGACCAATGAT 394
Qy      67 TTGCAATTCAGCTGGGACAGAGGTATCTATCCGCGCTTACAGCTGAAACAT 126
Db      395 TTTCAATGGAACAGATGTAAGTATGTTGCCACGTAATTAACGCAATTCGAT 454
Qy      127 GAAGCTGACGAA---GGGCCATCAGACAGTACGCGTTATTCGTAATGGGAGAA-- 181
Db      455 TTAGAAAAAGAAATGATGATTAACAATTCATATGCTTTTGTGGAATGGGAGCACT 514
Qy      182 -AACTGGAATATATATTTGGCCGAGCGCATTAAGAAAGCGCTTATCTGTGGGG 240
Db      515 GACATTAATCAACAGTCTCTATGATGGGTAACGAAAGGATTAATGGCCGATG 574
Qy      241 CTTTATTTTCCGGGCTATGCGAGTACGAAACAAATGATGCGGAATACCGTTCAAT 300
Db      575 CTTTACTATGCTACATTTGCGACTTATGCTGACCAACTTAAAGGACCAACAGGATC 634
Qy      301 GTCCCGATGAGTTGTGACATGGGTGCTGCTGATGCTGCTTGGAGAAGCTAA 360
Db      635 AATCCGCTGATGATATTTCTCAAGTTTGAAGATTTGTAATCTGATGATGTTAT 694
Qy      361 GAAAGATTCGATCTTTTACGATTTGAGAAATTAATGATCTATTTGATACAGTTT 420
Db      695 GAAATTTTACTTTTATATCAATTTATGATAGGCGCAATATATATCTTGGCTTGCACC 754
Qy      421 CCGCTTCACTGATATTTGACAGCGGACGCGCCGAAACCTGACGATAGAACCAAGACA 480
Db      755 CCACTTCACTATACATTTTACAGATGCTTGTGTAATGCAATGTTGTAAGACCGGATAA 814
Qy      481 GAGCGCTCAAACTGATGATTAATCAACCTGATGTAAGCAAGCGCGCGCATTA 540
Db      815 ACAGGCATTAACCATTCGAAAGATGTAAGGCTTGAAGAAATGCGCTGCGCTAA 874
Qy      541 TGGCATGTAACCAATCTGACGAAATATACAGAAATGACAGCGGAAATTAAGAGCAA 600
Db      875 TGGCATGACGAAATTTAAGGCTTATCATTTGTGTGTCACCAAAATCCGCGCAAGATTA 934
Qy      601 GAGATGGCGGATTAAGCGCTTCTGCTTTTGGCCAAAGCTTGAAGATGTTGTGCTGCG 660
Db      935 ATGATGGGAACTTGAATTTGACACCGTTTGGGCAAGGCGGCTTGAAGTATCA 994
Qy      661 GGGGATTAACACGCGCTTCCGCTTGTGACAGCTTTATTTGAAAGAAATCTAG 720
Db      995 GGTGATTTTACGCGCTGACAGCTTTTCTTGGGTAACATCTGGAAGAAATATATCTGAA 1054
Qy      721 CCGGCGCGGATTAAGAAAGATGTAACAGCGCTTTTCAATTTTGGCAATATGAC 780
Db      1055 AAAGCAAAATGAAACAGAGGCTTAACAATCTTGTTCATATCTTCTGTAAT 1114
Qy      781 ATACCAAGGCGCGAGTATACGAGAAAGCAAAATTCATTAACCAATATATCTTCC 840

```

Db 1115 ATCCAAAGGTGTGTTTGTGACAAATGAGGGGAAAAGGATTATACCTATACCTCA 1174  
Qy 841 GTGATGTGCAACAAATGGAATACTATTTCCACACATGACAATCGGAAATCCAA 900  
Db 1175 GCTATGTGCAACAAATGGAATACTATTTCCACACATGACAATCGGAAATCCAA 1234  
Qy 901 AAGTAATTTATTTTCAATGAAAGCTTGTGAGCTTAAAGTGTTCGGCTAA 960  
Db 1235 GCGGTTCCTTAATGCTGAAATTTAAATGCAAGATTATTAATTGAGTGGGAT 1294  
Qy 961 GCAGAGAGATTTCAATGAGCTTAAATGGA 993  
Db 1295 CGTAAACAGATTTAAAGCAATTAATCAAGTA 1327

RESULT 10  
US-10-282-122A-34300  
Sequence 34300, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Foreyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: Patent version 3.1  
SEQ ID NO 34300  
LENGTH: 978  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-10-282-122A-34300

Query Match 16.1%; Score 160.2; DB 7; Length 978;  
Best Local Similarity 49.9%; Pred. No. 4e-37;  
Matches 458; Conservative 0; Mismatches 453; Indels 6; Gaps 2;

Qy 4 ATGTCGACAGCTTACATGGAAGCTGACCGTAAAGATGATGAGAGACATG 63  
Db 1 ATGTGACGCTATTTCTTATATACAAACAGCTTATCATTTAGCTAGACATG 60

Qy 64 GATTTGCAATTCAGCTGGGGAAGAAGTATCTCTATCCGGCCGTTACAGCTGAAC 123  
Db 61 GACTTTCATTTGAAATTTATATGATATCCCAACCATTTGTCACAGTATATCACTACCA 120  
Qy 124 AATGAAGCTGACGGAAGGGCCATGACACAGATAGCGGTTATGCGGTATGGGAGAAA 183  
Db 121 TTTGATCTAGATTC--GACATGCGCTTGAATATGTTTGTGGAACAAATTTAAA 177  
Qy 184 CTTCGAAATATATTTATTTCCGACGCAATTAATGAAGCGGTTATCTTGTGGCGCTT 243  
Db 178 GTAGACGTTATAGATTTGTGATGATGATATTAAGAAAAGTTTGTATTTGAAACAT 237  
Qy 244 TATTTTCGGGCTATGCGGAGTACGAAAAACGATACGGAAGATACCGTTCAATTGC 303  
Db 238 TACTTCACTGAGGAGCCCTCATACAGTACCCATTAACGTTATGTTATTTAACTTAGCA 297  
Qy 304 CCGCATGATTTGTGACATGCGGTGCTGATGCTGATCTGATCTTTGGAACGTAAGAA 363  
Db 298 CCGAGAGTTTATTTGTTGGGTTTATGTTTAAATTAAGATTAAGCAATTTAAACAA 357  
Qy 364 AAGATTCGATCTTAAAGATGATGAGAAAAAATTAATGATCTATTTGATACAGTTTACG 423  
Db 358 AAGTTTAAAGATCAATTTATGATGATGATGATGATGATGATGATGATGATGATG 417  
Qy 424 CTTCAGTATATTTGTCAGACCGGACCGGCGCAATCTGACATGATGACCAAGAGAC 483  
Db 418 TTACATTTGATGCTCATGATGAAACAGGACATACCGTACATGACATGACATGAC 477  
Qy 484 GGCCTCAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 543  
Db 478 TTATTAATGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 537  
Qy 544 CATGTAACCAATCTGACGATATACAGGATCAGACCGAATTAAGAGCAAGAG 603  
Db 538 CATGTAACCAATCTGACGATATTAATTAATTAATTAATTAATTAATTAATTAATG 597  
Qy 604 ATGGCGGATTAAGCTTTCTGCTTTTGGCCAAAGCTTGAATCTGTGCTGCGGGG 663  
Db 598 ATAGGTAAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 657  
Qy 664 GATTAACACCGCTTCCGCTTTTGGTTCAGAGCTGTTATTTGAAACATCTAGAGCC 723  
Db 658 GGTATATCTTCAACAGACGTTTATGAGGCTCATATTTAAGACCAACATCACTGCT 717  
Qy 724 GCGCCGATGAAGCAAGGATGTAACGCGCTTTCATATTTGGCAATATGACAT 783  
Db 718 TCCATTAATGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 777  
Qy 784 CCAAGGCGCAGTGTAAACGAGAGCAAGATTCATTAACGATTAATCTTCCGT 843  
Db 778 CCGAAGGTGCAAGTAT--CGATGCCAATTAATTAATTAATTAATTAATTAATTAATG 834  
Qy 844 ATGTGCAAGAACTGGAATCTATTTCCACCATATGACATGAGGCAATTCAGAAA 903  
Db 835 ATGAAAGTAAAGAAAGGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 894  
Qy 904 GTAATTTATTTCAATG 920  
Db 895 ATTAATTTAATGTAAGA 911

RESULT 11  
US-10-724-972A-1455  
Sequence 1455, Application US/10724972A  
Publication No. US20040147734A1  
GENERAL INFORMATION:  
APPLICANT: Doucette-Stamm, Lynn  
APPLICANT: Bush, David  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PAT03-16  
CURRENT APPLICATION NUMBER: US/10/724,972A  
CURRENT FILING DATE: 2003-12-01

PRIOR APPLICATION NUMBER: 09/450,969  
 PRIOR FILING DATE: 1999-11-29  
 PRIOR APPLICATION NUMBER: 09/434,001  
 PRIOR FILING DATE: 1998-08-13  
 PRIOR APPLICATION NUMBER: 60/064,964  
 PRIOR FILING DATE: 1997-11-08  
 PRIOR APPLICATION NUMBER: 60/055,779  
 PRIOR FILING DATE: 1997-08-14  
 NUMBER OF SEQ ID NOS: 7544  
 SEQ ID NO: 1455  
 LENGTH: 996  
 TYPE: DNA  
 ORGANISM: S. epidermidis  
 US-10-724-972A-1455

Query Match 16.0%; Score 159; DB 7; Length 996;  
 Best Local Similarity 49.8%; Pred. No. 9, 2e-37;  
 Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

2 ATATGTGCAAGCTTACATTGGAAGTCTGACCGTAACATGTATTAGCAAGACA 61  
 14 ATATGTGACTGCGCATTTCTTATATACAAACAGTACCATTAATTAGTAGACACA 73  
 62 TGGATTTTGCATTTGAGCTGGGAGACAGAGTATCTATCCGCGCGTTACAGTGA 121  
 74 TGGACTTTGCAATTAATTTAAATGATCCCAACCATTTGTCACAGCAATATCACTAC 133  
 122 ACAGTGAAGCTGACGGAAGGCGCCATCAGACAGTACGCGTTATCGGTATGGGAGAA 181  
 134 AATTGATCTAGATTCACAGATGCTCTT--GAATATGTTTGTGGACAAATTTAA 190  
 182 AACTTGAATATATATTTATTTGCGGACGCGATTATGAAACGCGTTATCTGTGCGGCGC 241  
 191 AAGTAGGACGTTATGATTTGGTATGATTAAGAAAGAGTTAGCTATTTGAAAC 250  
 242 TTTATTTTCCGCGGCTATGCGAGTACGAAAAACATACGGAATACCTTCATCTG 301  
 251 ATATCTTACGTTGAGAGCTCTATACATACCATTAACGTTATGTTTAACTTAG 310  
 302 TCCCGCATGAGTTTGTGACATGGGTGCTGCTGCTGCTGCTTGGAGACGTAAG 361  
 311 CACCTGAGAGTTTATTTGTTGGTTTATGTTTAAATTAAGTATTTAGCAATTAAC 370  
 362 AAAAGATTGATCTTTAAAGATTTAGAGAAAAAATTAGATCTATTTGATACAGTTTAC 421  
 371 AAAAGTTAAGAAATCAATATTTATGATGAAAAAATACGACTTTGATATGCTCTC 430  
 422 CGCTTCACTGATATTTGTCAGACCGGAGCGGCGGAACTCTACATTAACCAAGACAG 481  
 431 CTTTACATTTCTATGCTGCTGATGAAACAGACATACCGTACCATTAACCTCAATG 490  
 482 ACGGCTCAAGTCTATGATATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541  
 491 GCTTATTAATGTTAAAGATTAATTTATGTTTACCTTAAACAAATGACCTTAATTAAGTT 550  
 542 GGCATGTACCAATCTGACGACATATACAGAAATCAGCCAGCAATTTAGAGCAAG 601  
 551 GGCATGTATCTAATCTTAAGAAATTAACGCTTTTAAACCCACAGAAATCAACATCAT 610  
 602 AGATGGGCGGATTAAGCCCTTTCTGCTTTTGGCCAAAGGTTAGAACTTGTCTGCGCG 661  
 611 TAAATGTTAAAGTCTAGTAAGATCAATGAGCTGTGAGACGGAACAATAGCTTACCGG 670  
 662 GGGATTTATACACCGCTTCCGCTTTCAGAGCTGTTTATTTGAAAGAACATCTAGAGC 721  
 671 GTGTTATAGTCAACAGATGTTTATAGCGCTACATATTTAAGACACCAATAGGCT 730  
 722 CCGGCGCGGATGAACAGAAAGTGTACAGCGCTTTTCAATTTTGGCAATATGACGA 781  
 731 GTTCCCATTAATGAAGATGAATAATTTAATGATTTGTTTAAAGTCTGAATCACTAGTA 790  
 782 TACCAAGGCGCGGTATTAACGAGAAAGCAAAATTCATTATAGCAATATCTTCCG 841

Db 791 TCCCTCAAGTGCAGTTAT---CGATGCCAATTAATACATTACACAAATATCAATTAG 847  
 Qy 842 TGAATGCAAGCAAGTGAAGTCTATTTTCCACCATATACATGCGGAAATCCAA 901  
 Db 848 TGAATGCAAGTGAAGTGAAGTCTATTTTAAATTAATGACCTTATTTGCAATCAAAATTTTCA 907  
 Qy 902 AAGTAAATTTATTTTCATGA 920  
 Db 908 AATTAATTTAATCTAGACA 926

# RESULT 12

US-10-282-122A-7960  
 Sequence 7960, Application US/10282122A  
 Publication No. US20040029129A1  
 GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zyckind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: EILTRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308

Remaining Prior Application data removed - See File Wrapper or PAM.  
 NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 7960

LENGTH: 993  
 TYPE: DNA

ORGANISM: Staphylococcus aureus  
 US-10-282-122A-7960

Query Match 14.7%; Score 146; DB 7; Length 993;  
 Best Local Similarity 47.5%; Pred. No. 7, 7e-33;  
 Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

4 ATGTGCAAGTCTTACATTGGAAGTCTGACCGTAACATGTATTAGCAAGACAATG 63  
 1 ATGTGCAAGATTTCAATATCAAACTTTAAATATCAATCACTTCTTGGACGACGATG 60  
 64 GATTTGCAATTTGAGCTGGGAGACAGAGTATCTATCCGCGCGCTTACAGTGAAC 123  
 61 GATTAATGATATCAATTAATGATGTTGCGCAGCAGTACGCTTGAATAATATATCTTGGACA 120

124 AGTGAAGCTGACGGAAGGCGCCATGACACAGTACCGCTTATCGGTATGGGAGAA 183  
 121 TCTCGACT--GGACGACAGGCGCCAAACGCAATATGCTTTATTTGTCAGAGAACAT 177  
 184 CTGGAAATATATATTTTGGCCGACGATTAAGAAAGCGTTTATCTTGTGGCGCTT 243  
 178 ATGGAAGGTTTATTTATGATGATGTTAATGAAATGCGGTCTATTTGACACAA 237  
 244 TATTTTCCGGCTATGCGGATGACAAAAACGATCGGGAGATCCGTTCACTTGC 303  
 238 TATTTCCGGGTTATGTTATATATGCTCAACACAAAGCGGACCGGATGATATTAG 297  
 304 CCGGATGATTTGTGACATGGGCTGTGACGTCTGCACTTTTGGAGACGTAAGAA 363  
 298 CAAATGAAATTTGTGACATGATTTGGATATACAAAGCATTAAGATATGAAACAA 357  
 364 AAGATTCGATCTTTAAGATTTAGAGAAAAATTAATCTATTGGATACAGTTTACG 423  
 358 CAGACATCCCAATATACATGTTAGAGGATATATTAAATGACATCGGTGAAGTTCCGCA 417  
 424 CTTCATGATATTTGTCACGCGGACGCGCCAAACCTGACATACCAAGACAGAC 483  
 418 TTGCAATTATCATGTTCCATGACATGCACTGACATTCAGTGAATTCATTAAAGGGG 477  
 484 GGCCCTCAAGCTATGATATATCACTGGGTGATGACGAAAGCGCCGCTTATATAG 543  
 478 GAAGTGCTTATTAAGATTAACCTTATGTTGTTTAAACAATATCATCAAGCTTAAGT 537  
 544 CATGTAAACATCTGACGATATACAGAAATGACGCAAGCAATTAAGACAAAGAG 603  
 538 CATTAATGATTAATTAAGACATATATCAATATTTCTTATCCACAAACAGCAAGTTA 597  
 604 ATGGCGGATTAACCTTCTTCTTGTGGCCAGGCTTAAAGAACTGTGGTCTGCGGGG 663  
 598 TTGGAAGGTGTAAAGATTAACCTTATGAGCAATGAAAGAGTACATTTGATGCGAGG 657  
 664 GATTATACACCGCTTCCGCTTGTGACAGCTGTTTATTTAAAGAAACATCTAGAGCG 723  
 658 GATTTACTTCACTAGAGCTTGTGAGATATGCAATTTATGAAACCAATCTGCTCA 717  
 724 GCGGCGGATGAAGCAAGGTGTACAGCGCTTTTCAATTTTGGCAATATAGAGATA 783  
 718 AACATGATTAAGAAATGATTTAATGAATGATTTATTTAATGATGCGGTAATATA 777  
 784 CCAAGGCGGCTGATAGGAGAAAGCAATTCATTAATGCAATTAATCTTCTG 843  
 778 CCGATTTGATTTGATGCTCGCATGATGCTGACATCACTATGATGATGACAGCGTA 837  
 844 ATGTGCAAGCAACGTAAGCTATTTTCCACCACTATGACATGCGCAATCCAAATA 903  
 838 ATTAATTTAATCAAGAACTTATATTAATTAATGATTAATGCAATGAATATAGTA 897  
 904 GTAAATTTATTTATGAGAACTGATGCTTTTGGAGCTTAAGTGTTCGCTTAAGCA 963  
 898 TTAAGCTCAGATGATTTAATTAATGAAGAAATGATGATTTTAAGCTGAGAG 957  
 964 GAAGAGATTTATGATGCTTAATTA 989  
 958 CATATCATATTTAAGAAAGTTGAATGA 983

RESULT 13  
 US-08-781-986A-444/C  
 ; Sequence 444, Application US/08781986A  
 ; Publication No. US2003005443A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 5255  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville

STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/781,986A  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benson, Bob  
 REGISTRATION NUMBER: 30,446  
 REFERENCE/DOCKET NUMBER: PB248BP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 444:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1146 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-781-986A-444  
 Query Match 14.7%; Score 146; DB 2; Length 11466;  
 Best Local Similarity 47.5%; Pred. No. 2,9e-32;  
 Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

4 ATGTGCAAGCTTATCATTTGAAATCTGTGACCGTTAAACATGATTAAGCAAGAAATG 63  
 3688 ATGTGCAAGCTTATCATTTGAAATCTGTGACCGTTAAACATGATTAAGCAAGAAATG 3629  
 64 GATTTGCAATTCAGCTGGGGAACAGAGGATTTCTATCCGCGCTTACAGCTGAAAC 123  
 3628 GATTTATGATTAATCATTAATGATGTTCCGACAGATGACCGCTTAATGATGTTGAA 3569  
 124 AGTGAAGCTGACGGAAGGCGCCATGACACAGTACGCTTTATCGGTATGGGAGAAA 183  
 3568 TCTTGCACT--GGACGACAGGCGCCAAACGCAATATGCTTTATGCGACAGAAAGAT 3512  
 184 CTGGAAATATATTTTGGCCGACGATTAAGAAAGCGTTTATCTTGTGGCGCTT 243  
 3511 ATGGAAGGTTTATTTATGATGATGATGTTAATGAACATGCGCTTCAATTAACAA 3452  
 244 TATTTTCCGGGCTATGCGGAGTACGAAAAACGATACGGGAAAGATCCGTTCACTTGT 303  
 3451 TATTTCCGAGTTATATGTTCAATGATGATCAACACAAAGCGGACCGATTAATTAAG 3392  
 304 CCGCATGATTTGTGACATGGGTGCTGTGACGTCTGCACTTTTGAAGACGTAAGAA 363  
 3391 CAAATGAAATTTGTGACATGATTTTGGATATACAAAGCATTAAGATGAAACAA 3332  
 364 AAGATTCGATCTTTAAGATTTGAGAAAAATTTGATCTATTTGATACAGTTTAACG 423  
 3331 CAGCATCCCAATATACATGTTGATGCTGTATTTAATGAACATGCGTGAAGTTCCGCA 3272  
 424 CTTCATGATATTTGTGACGCGGACGCGCCAAACCTGACATGAAACCAAGACAGAC 483  
 3271 TTGCAATTATCATGTTCCATGATGACATGACATACAGTGAATTTCAATTAAGAGGT 3212  
 484 GGCCCTCAAGCTATGATATATCACTGGGTGATGACGAAACGCCCGCATTTATATG 543  
 3211 GAAGTGCTTATTAAGATTAATCTTATGTTGTTTAAACAATCATCCAGCTTAATTTG 3152  
 544 CATGTAAACATCTGACGATATACAGAAATGACGCAAGCAATTAAGACAAAGAG 603  
 3151 CATTAATGATTAATTAAGCAATATATCAATATTTCTTATCAAGCAAGCAATTTTA 3092



QY 604 ATGGCGGATTAAGCCCTTCTGCTTTGGCCAGGCTTAGAAGCTGTGCTGCGGG 663  
 DB 3091 TTGGAAAGGTGAAGATTGAACCTTTAGGCAATAGAGGATCATTTGGATCCAGGT 3032  
 QY 664 GATTATACACGCGCTTCCGGTTGTGACAGCTGTTTATTGAAGAACAATCTAGACCG 723  
 DB 3031 GGATTTCTTCAACTGAGCGCTTTGTGAAATGCGATTATGAAGAACAATCTGCA 2972  
 QY 724 GCGGCGGATGAAGAAAGGTGTAAACGCGCTTTTCAAAATTTGGCAAAATATGACGAT 783  
 DB 2971 AACATGATTAAGAAATGATTTAATGAATGCAATTTATTATAGATGCGTAAATTA 2912  
 QY 784 CCAAGGCGGAGATGAAGAAAGGAAAGCAAAATCTATTAACGCAATATCTCCGTG 843  
 DB 2911 CCGATTGAATTTGACGTCGCGATGATGTCGACATCACTAATGATGATCAGACCGTA 2852  
 QY 844 ATGTGCAAGAACTGGAACCTACTATTTCCACACTATGACATGCGCAATCCAAA 903  
 DB 2851 ATAAATTTACTACAGAACGTTATATTAATTAATGCGCAATGAATTAAGTGA 2792  
 QY 904 GTAAATTTATTCATGAAGACCTTGACTTTGGAGCCTAAAGTGTTCGCTAAAGCA 963  
 DB 2791 TTAAGCTCAGATGATTTAATTAATGAAGAAAGATATGACATTTTAAAGCTGAGAG 2732  
 QY 964 GAAGAGATTTCTAGAGCTTAATTA 989  
 DB 2731 CATATCACTATTGAAGATTGAATGA 2706

## RESULT 14

US-10-329-624-444/C  
 Sequence 444, Application US/10329624  
 Publication No. US20040043037A1

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
 Gil H. Choi

Patrick S. Dillon  
 Craig A. Rosen

Steven C. Barash  
 Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/329,624

FILING DATE: 27-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/956,171

FILING DATE: October 20, 1997

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248PIDI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 444:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 11466 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 444:  
 US-10-329-624-444

Query Match 14.7%; Score 146; DB 7; Length 11466;  
 Best Local Similarity 47.5%; Pred. No. 2.9e-32;  
 Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

QY 4 ATGTGCAAGCTTATTCATTGGAAGTGTGACCGTAAACATGATTTAGCAAGCAATG 63  
 DB 3688 ATGTGCAAGCTTATTCATTGGAAGTGTGACCGTAAACATGATTTAGCAAGCAATG 3629  
 QY 64 GATTATTAATTCAGTGGGACAGAGGATTTCTATTCGCGCGCTTACAGCTGAC 123  
 DB 3628 GATTATTAATTCAGTGGGACAGAGGATTTCTATTCGCGCGCTTACAGCTGAC 3569  
 QY 124 AGTGAAGCTGACGGAAGGCGCCATCAGACAGTACGCGTTATCGGTATGGGAGAAA 183  
 DB 3568 TCTTGCACT--GCAACGACAGGCGCAAGCAATATGCTTTATGGCAGAGAAAGAT 3512  
 QY 184 CTGGAATATATATTTTCCGACGCGCATTAATGAAGCGGTTATCTTGCGCGCTT 243  
 DB 3511 ATGGAAGCTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3452  
 QY 244 TATTTCCCGGCGTATGCGAGATGCAAAAAAGATACGGGAATATCCGTTCAATTTG 303  
 DB 3451 TATTTCCCGGCGTATGCGAGATGCAAAAAAGATACGGGAATATCCGTTCAATTTG 3392  
 QY 304 CCGCATGAGTTTGTGACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
 DB 3391 CAAATGAATTTGTGACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3332  
 QY 364 AAGATTCATCTTTAAGATGATGAGAAAAATTTAGATCTATTGATACAGTTTAA 423  
 DB 3331 CAAGCATCCCAATATCATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3272  
 QY 424 CTTCAGTATATTTGACAGCGGACGCGGCGGCAACCTGACATGATGACAGAGAC 483  
 DB 3271 TTGCATATCATGTTTCCGATGCACTGACATGACATGACATGACATGACATGACAT 3212  
 QY 484 GGCCTCAAGTCTATGATATCACTGATGATGATGATGATGATGATGATGATGATGATG 543  
 DB 3211 GAAGTGTATTAAGATATCTATGATGATGATGATGATGATGATGATGATGATGATG 3152  
 QY 544 CATGTAACCAATCTGACAGATATTAAGAAATCAGACGAGCAATTTAGAGCAAGAG 603  
 DB 3151 CATTAATGAATTTAAGACATATTAATTAATTTCTCTTATCCAGCAACAGCAATTTA 3092  
 QY 604 ATGGCGGATTAAGCCCTTCTGCTTTGGCCAGGCTTAGAAGCTGTGCTGCGGG 663  
 DB 3091 TTGGAAAGGTGAAGATTGAACCTTTAGGCAATAGAGGATCATTTGGATCCAGGT 3032  
 QY 664 GATTATACACGCGCTTCCGGTTGTGACAGCTGTTTATTGAAGAACAATCTAGACCG 723  
 DB 3031 GGATTTCTTCAACTGAGCGCTTTGTGAAATGCGATTATGAAGAACAATCTGCA 2972  
 QY 724 GCGGCGGATGAAGAAAGGTGTAAACGCGCTTTTCAAAATTTGGCAAAATATGACGAT 783  
 DB 2971 AACATGATTAAGAAATGATTTAATGAATGCAATTTATTATAGATGCGTAAATTA 2912  
 QY 784 CCAAGGCGGAGATGAAGAAAGGAAAGCAAAATCTATTAACGCAATATCTCCGTG 843  
 DB 2911 CCGATTGAATTTGACGTCGCGATGATGTCGACATCACTAATGATGATCAGACCGTA 2852  
 QY 844 ATGTGCAAGAACTGGAACCTACTATTTCCACACTATGACATGCGCAATCCAAA 903  
 DB 2851 ATAAATTTACTACAGAACGTTATATTAATTAATGCGCAATGAATTAAGTGA 2792  
 QY 904 GTAAATTTATTCATGAAGACCTTGACTTTGAGAGCTTAAAGTGTTCGCTAAAGCA 963

Db 2791 TTTAAAGTCAAGATGATTTTAAATAGAAAAGATATGACGATTTTAAAGCTGAGAG 2732  
 QY 964 GAAGAGAGTTTCATGAGCTTAATTA 989  
 Db 2731 CATATCACTATTAGAAAGTTGAATGA 2706

## RESULT 15

US-09-815-242-8383  
 ; Sequence 8383, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zykkind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8383  
 ; LENGTH: 1002  
 ; TYPE: DNA  
 ; ORGANISM: Staphylococcus aureus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(1002)  
 ; US-09-815-242-8383

Query Match 14.3%; Score 142.8; DB 3; Length 1002;  
 Best Local Similarity 47.3%; Pred. No. 7.2e-32;  
 Matches 466; Conservative 0; Mismatches 517; Indels 3; Gaps 1;

QY 4 ATGAGCAAGCTTACATTTGGAACCTGTCGCTAAACATGTTAGCAAGAACATG 63  
 Db 10 ATGAGCAAGCTTACATTTGGAACCTGTCGCTAAACATGTTAGCAAGAACATG 69  
 QY 64 GATTTGATTTCACTGCGGAGCAGAGGTGATTTCTATCCGCGCGCTTACAGCTGAAC 123  
 Db 70 GATTTGATTTCACTGCGGAGCAGAGGTGATTTCTATCCGCGCGCTTACAGCTGAAC 129  
 QY 124 AGTGAAGCTGACGGAAGGCCCATGACAGACAGTACGCGTTATCGGTATGGGAGAAA 183  
 Db 130 TCTCGACT--GGACAGCAGAGCCAAAGCAATATGGCTTATGGTACAGAAACAGAT 186  
 QY 184 CTGGAATATATTTATTTGCGCAGCGCATTTAATGAAGCGTTTATCTTGTGCGCGCTT 243  
 Db 187 ATGGAAGCTTTATTTATGATGATGCTTATTAATGAACATGCGCTTCAATTAACACAA 246  
 QY 244 TATTTCCGGCTATGCGAGTACGAAAAACGATACGGGAAGATACCGTTCACTGTC 303

Db 247 TATTTCCGAGTTATGATTCATATGCTCAACACACAAAGCGGATGCAATTAATACG 306  
 QY 304 CCGCATGAGTTTGTGACATGCGGTGCTGCAATCTGTCACTTTGGAAAGCTTAAGAA 363  
 Db 307 GAAATGAATTTGTGACATGATTTTGGATATACAAAGCATTTGAAGTATGAACAA 366  
 QY 364 AGATTCGATCTTTAAGATTTGAGAAAAAATTAATGATCTATGATACAGTTTAACG 423  
 Db 367 CAAGCATCCCAATATCATGTTGTAGCTGTATATTAATGATATGCGTGAAGTTCCGCCA 426  
 QY 424 CTTCATCTGATATTTGACAGCCGAGCGGCCAAACCTGACGATPAGAACCAAGACGAC 483  
 Db 427 TTGCATATATCATGTTCCGATGCACTGACATTCAGTCAATTTCAATTTAAGAGGT 486  
 QY 484 GGCCTCAAGTCTATGATTAATCAACCTGCTGTCATGACGAAACGCCGACTTTATATG 543  
 Db 487 GAAGTATATTAAGATTAATCACTATTTGCTTAAACAAATCATCAAGCTTATGATGG 546  
 QY 544 CATGTAACTATCTGACGAAATATACAGAAATCAACCGAAGCAATTAAGAGCAAGAG 603  
 Db 547 CATATATGATTTAAGACAAATATATCAATATTTCTTATCCAGCAACGCAAAATTA 606  
 QY 604 ATGGCGGATTAAGCCCTTTCTGCTTTGGCCAGCTTGAAGCTGTGTGCTGCCGGGG 663  
 Db 607 TTGGAAGTGAAGCAATGAACCTTTAGGCAATGAAGCAAGTACATTTGATTTGCCAGGT 666  
 QY 664 GATTAATACCGGCTTCCGCTTGTGACAGCTGTTTATTTGAAAGCAATCTAGAGCCG 723  
 Db 667 GATTTACTTCAATGAGCGCTTTGTGAGAAATGACATTTATGAAGCAATTTGCTCA 726  
 QY 724 GCGGCGGATGAAGCAAGGTGTAACAGCCGCTTTCAATTTGGCAAAATATGACATA 783  
 Db 727 AACATGATTAAGAAATGATTAATGAATGATTTTATTAATGAATGCGGTAAATATA 786  
 QY 784 CCAAGGCGGATGATTAAGGAAAGAAAGAAATCATTAATCGCAATTAATCTCGGTG 843  
 Db 787 CCGATGGAATTTGATCGTCCGATGATGCTGACATTAATCAATATGATATGACAGCGTA 846  
 QY 844 ATGCAAGCAAACTGGAATCTATTTCCACCATATGACATCGGCAATCCAAATA 903  
 Db 847 ATTAATTTAATCAAGAAAGCTTATATTAATGATTAATGACAGCAATGAATGATGACA 906  
 QY 904 GTAATTTATTTCAATGAAGACCTTGAAGCTTGTGAGCTTAAAGCTTTTCCGCTAAACA 963  
 Db 907 TTAAGCTCAAGATGATTTAATTAATGAAGAAATGATGACGATTTTAAGCTGAGAG 966  
 QY 964 GAAGAGTATTTCAATGAGCTTAATTA 989  
 Db 967 CATATCACTATTAAAAAGTTGAATGA 992

Search completed: March 18, 2006, 13:28:30  
 Job time : 998 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using ew model

Run on: March 18, 2006, 11:54:08 ; Search time 452 Seconds

(without alignments)  
5138.006 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996  
Sequence: 1 catatgacacacagctctac.....atgagcttaataagatcc 996

Scoring table:  
IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA-New\*

1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	16.0	990	US-10-793-626-413	Sequence 413, App
2	159	16.0	2975	US-10-793-626-345	Sequence 345, App
3	159	16.0	4114	US-10-793-626-4184	Sequence 4184, App
4	93	9.3	975	US-11-074-176-367	Sequence 367, App
5	83.8	8.4	975	US-11-074-176-369	Sequence 369, App
6	38.8	3.9	720	US-10-750-185-59863	Sequence 59863, A
7	38.8	3.9	720	US-10-750-623-59863	Sequence 59863, A
8	36.2	3.6	1272	US-11-096-568A-29311	Sequence 29311, A
9	35.8	3.6	583	US-09-925-065A-392869	Sequence 392869, A
10	35.4	3.6	419	US-09-925-065A-654182	Sequence 654182, A
11	35.4	3.6	419	US-09-925-065A-654183	Sequence 654183, A
12	35.4	3.6	5901	US-10-932-182A-2969	Sequence 2969, App
13	35.4	3.6	5901	US-10-932-182A-2969	Sequence 2969, App
14	35	3.5	961	US-11-072-175-102	Sequence 102, App
15	34.6	3.5	583	US-09-925-065A-392868	Sequence 392868, App
16	34.4	3.5	4429	US-10-753-195-9	Sequence 9, App11
17	34.2	3.4	583	US-09-925-065A-392870	Sequence 392870, A
18	34.2	3.4	1835	US-09-925-065A-45228	Sequence 45228, A
19	34.2	3.4	1835	US-09-925-065A-45229	Sequence 45229, A
20	34.2	3.4	1835	US-09-925-065A-45230	Sequence 45230, A

21	34	3.4	601	US-09-925-065A-348226	Sequence 348226, App1
22	34	3.4	193084	US-11-121-086-82	Sequence 82, App1
23	33.8	3.4	670	US-09-925-065A-556802	Sequence 556802, A
24	33.6	3.4	523	US-09-925-065A-155341	Sequence 155341, A
25	33.6	3.4	556	US-09-925-065A-77072	Sequence 77072, A
26	33.4	3.4	572	US-09-925-065A-29870	Sequence 29870, A
27	33.4	3.4	580	US-09-925-065A-321496	Sequence 321496, A
28	33.4	3.4	653	US-09-925-065A-96912	Sequence 96912, A
29	33.4	3.4	659	US-09-925-065A-510272	Sequence 510272, A
30	33.4	3.4	958	US-09-925-065A-713894	Sequence 713894, A
31	33.4	3.4	958	US-09-925-065A-713895	Sequence 713895, A
32	33.4	3.4	1197	US-10-750-185-62875	Sequence 62875, A
33	33.4	3.4	1197	US-10-750-623-68875	Sequence 68875, A
34	33.4	3.4	1217	US-09-925-065A-686875	Sequence 686875, A
35	33.4	3.4	2830	US-09-925-065A-669245	Sequence 669245, A
36	33.4	3.4	2830	US-09-925-065A-669246	Sequence 669246, A
37	33.2	3.3	629	US-09-925-065A-346097	Sequence 346097, A
38	33.2	3.3	645	US-09-925-065A-860511	Sequence 860511, A
39	33.2	3.3	661	US-09-925-065A-690516	Sequence 690516, A
40	33.2	3.3	2037	US-10-932-182A-1230	Sequence 1230, App
41	33.2	3.3	2037	US-10-932-182A-1230	Sequence 1230, App
42	33.2	3.3	156297	US-11-121-086-65	Sequence 65, App1
43	33	3.3	393	US-09-925-065A-148412	Sequence 148412, A
44	33	3.3	499	US-09-925-065A-845854	Sequence 845854, A
45	33	3.3	566	US-09-925-065A-828854	Sequence 828854, A

## ALIGNMENTS

RESULT 1  
US-10-793-626-413  
Sequence 413, Application US/10793626  
Publication No. US20050255478A1  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/10/793, 626  
CURRENT FILING DATE: 2004-03-04  
PRIOR APPLICATION NUMBER: 60/164, 258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 413  
LENGTH: 990  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-413

Query Match 16.0%; Score 159; DB 8; Length 990;  
Best Local Similarity 49.8%; Pred. No. 2e-35;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;  
2 ATATGTCACAGCTTACATGGAACCTGCGGTAACATGTAATTCAGCAAGCA 61  
8 ATATGTCACAGCTTACATGGAACCTGCGGTAACATGTAATTCAGCAAGCA 67  
62 TGAATTTGCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121  
68 TGAATTTGCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 127  
122 AATTTGCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181  
128 AATTTGCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 184  
182 AATTTGCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241  
185 AATTTGCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 244

QY 242 TTTATTTCCGGCTATGCGAGTACGAAAAAAGATACGGGAAGATCCGTTCACTTG 301  
 DB 245 ATTACTTCACTGCTGAGAGCTCATACAGTACCCATTAACCTTTATTTTAACTTGA 304  
 QY 302 TCCCGCATGAGTTTGACATGGGCTGTCACTCTGTCACTTTTGGAAAGCTTAAAG 361  
 DB 305 CACCTAGAGGATTATGTTGGTTTATGTTTAAAGATTAAGATTAAGCAATTTAAAC 364  
 QY 362 AAAAGATTCATCTTTTACGATTTGAGAAAAAATTAAGATCTATTGATACGTTTAC 421  
 DB 365 AAAAGGTTAAGAAATCAATTAATTAATGAATGAATGAATGAATGAATGAATGAAT 424  
 QY 422 CGCTTCACTGATATTGTCAAGCGAGCGGCGGAAACCTGACGATGAACCAAGACAG 481  
 DB 425 CTTTACATTTCACTGATGATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 484  
 QY 482 AGCGCTCAAAAGCTATGATTAATCAACCTGTCTCATGACGAACAGCCCGACTTTAT 541  
 DB 485 GCTTATTAATAGTTAAAGATTAATTAATGTTTCACTTAACAAATGAACCTTAATTA 544  
 QY 542 GGCATGTAACCATCTGACGAAATATACAGAAATCAACGGAAGCAATTAAGACGAAG 601  
 DB 545 GGCATCTATCTAATCTTAAGAAATTAACGCTTTTAAAGCAACAAATCAACATCAAT 604  
 QY 602 AGATGGCGGATTAAGCCCTTTCTGCTTTGGCCAAAGCTTAGAACTGTGTCTGCGCG 661  
 DB 605 TAATAGTAAAGTCTAGTAAATCAATGAGCTGTGAGCAAGAAATAGCTTACCGG 664  
 QY 662 GGGATTAATACACCGCTTCCCGGTTTGTCAAGCTGTTTATTTGAAAGAACTATGAGC 721  
 DB 665 GTGGTATATCGTCAACAGATCGTTTAAAGCGCTCAATTAATTAAGCAACCACTAGCT 724  
 QY 722 CGCGCGCGCATGAACGAAGTGTACAGCGCTTTTCAATTTTGGCAAAATATGAGCA 781  
 DB 725 GTTCCATTAATGAAGATGAATTAATTAATGATGTTTAAAGTCTTAAGTCACTGCTA 784  
 QY 782 TACCAAGGCGCAGTATTAACGAAAGCAAAATTCATTTATACCAATATATCTTCCG 841  
 DB 785 TCCTCAAGGTGAGTAT---CGATGCCAATAAATTAATTAATTAATTAATTAATTA 841  
 QY 842 TGATGCAACGAAGTGAAGTAACTATTTTCCACCACTATGACATGCGCAATCCAA 901  
 DB 842 TGATGCAAGTAAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 901  
 QY 902 AAGTAATTTATTCATGA 920  
 DB 902 AAATAAATTAATGAAGA 920

## RESULT 2

US-10-793-626-3345/C  
 ; Sequence 3345, Application US/10793626  
 ; Publication No. US20050255478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMMERY, WILLIAM JOHN  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 ; FILE REFERENCE: P03480US  
 ; CURRENT FILING DATE: 2004-03-04  
 ; PRIOR APPLICATION NUMBER: 60/164,258  
 ; PRIOR FILING DATE: 1999-11-09  
 ; NUMBER OF SEQ ID NOS: 4472  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3345  
 ; LENGTH: 2975  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 ; OTHER INFORMATION: nucleic acid sequence  
 US-10-793-626-3345

Query Match 16.0%; Score 159; DB 8; Length 2975;

Best Local Similarity 49.8%; Pred. No. 3,4e-35; Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTCACAAGCTTATATATGAAACGTCGACCGCTTAAACATGATTAAGCAAGAA 61  
 DB 1755 ATATGTCACCTGATTTCTTTATATTAACAAACAGCTTACATTAATTAAGTAA 1696  
 QY 62 TGGATTTGTCATTTGAGCTGCGGACAGAGTATCTTATCCGCGCTTACAGCTGA 121  
 DB 1695 TGGATTTGTCATTTGAGTATTAATGATATCCAAACATTTGTCAGCCATTAATCACTAC 1636  
 QY 122 ACAGTAAGCTGACGGAAGGCGCATCAGACACATGACCGCTTATCGGTATGGAGAG 181  
 DB 1635 AATTGATCTAATTCAGACATGCGCTT---GAATATGTTTGTGAAACAAATTTAA 1579  
 QY 182 AACTGGAATATATATTTATTTGCGGACGCAATTAATGAACGCTTATCTGCGGCGC 241  
 DB 1578 AAGTGAAGCTTATATATTTGATGATGATTAACGAAAGTTTATGATTTCCAGC 1519  
 QY 242 TTTATTTCCGGCTATGCGGAGTACGAAAAAAGATACGGAAGATACCGTTCACTTG 301  
 DB 1518 ATTACTTCACTGCTGAGAGCTCATACAGTACCAATTAAGTTATGTTTAACTTAG 1459  
 QY 302 TCCCGCATGAGTTTGACATGGGCTGTCACTCTGTCACTTTTGGAAAGCTTAAAG 361  
 DB 1458 CACCTAGAGGATTATGTTGGTTTATGTTTATTAATTAAGTATTAAGCAATTAAC 1399  
 QY 362 AAAAGATTCATCTTTAAGATTTGTAAGAAAAAATTAATCATGATTAAGTATCACTTTAC 421  
 DB 1398 AAAAGGTTAAGAAATCAATTAATTAATGAATGAATGAATGAATGAATGAATGAAT 1339  
 QY 422 CGCTTCACTGATATTGTCAAGCGGACGCGGAAACCTGACGATGAACCAAGACAG 481  
 DB 1338 CTTTACATTTCACTGATGATGAAGACGATACCGTACGATGAACCTCAACATG 1279  
 QY 482 AGCGCTCAAAAGCTTATGATTAATCAACCTGTCTCATGACGAACAGCCCGCACTTATAT 541  
 DB 1278 GCTTATTAATAGTTAAAGATTAATTAATGTTTCACTTAACCAATGAACCTTAATTA 1219  
 QY 542 GGCATGTAACCATCTGACGAAATATACAGAAATCAACGAAATTAAGAGCAAG 601  
 DB 1218 GGCATCTATCTAATCTTAAGAAATTAACGCTTTTAAAGCAAGAAATCAACATCAAT 1159  
 QY 602 AGATGGCGGATTAAGCCCTTTCGCTTTTGGCCAAAGCTTAGAACTGTGTGCTGCGG 661  
 DB 1158 TAATAGTAAAGTGTAGTAAATCAATGCGCTGTAAGAGCAAGAAATAGGCTTACCGG 1099  
 QY 662 GGCATTAATACACCGCTTCCCGGTTTGTCAAGCTGTTTATTTGAAAGAACTATGAGC 721  
 DB 1098 GTGGTATATCGTCAACAGATCGTTTATACGCGCTCAATTAATTAAGCAACCACTAGCT 1039  
 QY 722 CGCGCGCGCATGAAGAAAGTGTAAACAGCGCTTTTCAATTTTGGCAATATATGAGCA 781  
 DB 1038 GTTCCATTAATGAAGATGAATTAATTAATGATGTTTAAAGTCTTAAGTCACTGCTA 979  
 QY 782 TACCAAGGCGCAGTATTAACGGAAGACGAAATTCATTTATACGCAATATATCTTCCG 841  
 DB 978 TCCTCAAGGTGAGTAT---CGATGCCAATAAATTAATTAATTAATTAATTAATTA 922  
 QY 842 TGATGCAACGAAGTGAAGTAACTATTTTCCACCACTATGACATGCGCAATCCAA 901  
 DB 921 TGATGCAAGTAAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 862  
 QY 902 AAGTAATTTATTCATGA 920  
 DB 861 AAATAAATTAATGAAGA 843

## RESULT 3

US-10-793-626-4184  
 ; Sequence 4184, Application US/10793626  
 ; Publication No. US20050255478A1  
 ; GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN  
 TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 FILE REFERENCE: PUS480US  
 CURRENT APPLICATION NUMBER: US/10/793,626  
 CURRENT FILING DATE: 2004-03-04  
 PRIOR APPLICATION NUMBER: 60/164,258  
 PRIOR FILING DATE: 1999-11-09  
 NUMBER OF SEQ ID NOS: 4472  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 4184  
 LENGTH: 4114  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 OTHER INFORMATION: nucleic acid sequence  
 US-10-793-626-4184

Query Match 16.0%; Score 159; DB 8; Length 4114;  
 Best Local Similarity 49.8%; Pred. No. 4e-35;  
 Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTCACAAAGTCTTACATGGAAACGTGACCGTAAACATGATATGCAAGACAA 61  
 DB 1626 ATATGTCACCTGCCATTTCTTATATACAAAACACGTTACCATTTATGCTAGAACAA 1685  
 QY 62 TGGATTTTGCATTTGAGCTGGGGACAGAGGTGATTTCTATCCGCGCGCTTACAGCTGGA 121  
 DB 1686 TGGACTTTGCATTTGAAATTTAATGATCCCAACCATTTGTCACGCCATTTATCACTACC 1745  
 QY 122 ACAGTGAAGCTGACGGAAGGCCCATGACACAGTACCGCTTTATCGGTATGGGGAGAA 181  
 DB 1746 AATTGATCTTACATTCAGACATGCGCTT---GAATATGTTTGTGGAAACAAATTTAA 1802  
 QY 182 AACTTGAATATATATTTATGCGACGCGATTAATGAACGGTTATCTTGCGCGCC 241  
 DB 1803 AAGTGAAGCTTATGATTTGTTGATGATTAACGAAAGGTTAGCTTATTTGGAACC 1862  
 QY 242 TTTATTTTCCGGGCTATGCGAGTACGAAACGATACGGAAGTACCGTTCACTTG 301  
 DB 1863 ATTACTTCACTGCTGAGGCTCATACAGTACCCATTAAGTTATTTTAACTTAG 1922  
 QY 302 TCCCGCATGAGTTTGTGACATGGGTGCTGCTGCTGCTGCTGCTTGGAAAGCTTAAAG 361  
 DB 1923 CACCTGAGGAGTTTATGTTTGGTTTATGTTTATTAATGAATATGCGAATTAAC 1982  
 QY 362 AAAAGATTCGATCTTAACGATTTGAGAAATTAATGATCTATTTGATACAGTTTAC 421  
 DB 1983 AAAAGCTTAAAGAAATCAATATTAATGAATGAAGAAATTAACGATTTGATATCTCTC 2042  
 QY 422 CGCTTCACTGATATTTGTACAGCCGAGCGGCCGAAACCTGACATAGAACCAAGAGCAG 481  
 DB 2043 CTTTACATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2102  
 QY 482 ACGGCTCAAAAGTCTATGATTAATCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541  
 DB 2103 GCTTATATATGTTTAAAGTATATATGTTTACCTTAAACCAATGAACCTTAATAGATT 2162  
 QY 542 GGCATGTAAACAATGTGACGAATATACAGAAATGACACGGAACAAATTAAGAGCAAG 601  
 DB 2163 GGCATCTATCTTAATTAAGAAATTAAGCTTTTAAACGCAAGAAATCAACCAATCAAT 2222  
 QY 602 AGATGGCGGATTAAGCCCTTCTGCTTTGCGCAAGGCTTGAAGCTTGGCTGCGCG 661  
 DB 2223 TAAATGTAAGTGTAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2282  
 QY 662 GGAATATACACCGCTTCCGGGTTTGTACAGCTGTTTATTTGAAAGAACATCTAGAC 721  
 DB 2283 GTGTATATACGTCAACAGATGCTTTTATACGGGCTACATATTTAAGAACACCAATACCT 2342  
 QY 722 CGGGCGCGGATGAACGAAGGTGAACAGCGCTTTCAAAATTTGGCAATATATAGCA 781  
 DB 2343 GTTCCATATATGAATGAATTAATTAAGAAATGTTTAAAGTTCTAATAATCAGTACGTA 2402

QY 782 TACCAAGGCGCGAGTATAGCAAGAGAAAGAAATTCATTAAGCAATATACCTCCG 841  
 DB 2403 TCCCTCAAGGTGACGTTAN---CGATGCCAATTAATATATACACATATTCATTTG 2459  
 QY 842 TGAATGCAAGAAATGGAACACTATTTCCACACATATGACAAATGCGCAATCCAA 901  
 DB 2460 TGAATGAAAGTAAAGAAAGAAAGTATTAATTAAGCCTTACTTACCAATCAAAATTTCA 2519  
 QY 902 AAGTAAATTAATTTATGATGA 920  
 DB 2520 AATTAATTAATTAATGAAG 2538

RESULT 4  
 US-11-074-176-367  
 Sequence 367, Application US/11074176  
 Publication No. US20050250135A1  
 GENERAL INFORMATION:  
 APPLICANT: Kleenhamer, Todd R.  
 APPLICANT: Russell, William M.  
 APPLICANT: Altermann, Eric  
 APPLICANT: McAniff, Olivia  
 APPLICANT: Peril, Andrea Azcarate  
 TITLE OF INVENTION: Nucleic Acid Sequences Encoding  
 TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore  
 FILE REFERENCE: 5051-694  
 CURRENT APPLICATION NUMBER: US/11/074,176  
 PRIOR FILING DATE: 2005-03-07  
 PRIOR APPLICATION NUMBER: 60/551,161  
 PRIOR FILING DATE: 2004-03-08  
 NUMBER OF SEQ ID NOS: 381  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 367  
 LENGTH: 975  
 TYPE: DNA  
 ORGANISM: Lactobacillus acidophilus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)...(975)  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (0)...(0)  
 OTHER INFORMATION: ORF 892; bsba  
 US-11-074-176-367

Query Match 9.3%; Score 93; DB 12; Length 975;  
 Best Local Similarity 46.4%; Pred. No. 2e-16;  
 Matches 340; Conservative 0; Mismatches 390; Indels 3; Gaps 1;

QY 197 TATTGGCGAGGAGTAAATGAAGCGTTATCTTGCGCGCCTTATTTCCGGCT 256  
 DB 188 TATATTTGAGCGCTACAAATGAATGAAGTTTATGATGCGGAGCTCAATATCCAGAA 247  
 QY 257 ATGCGAGTACGAAATTAAGTATGAGGAGTACCGTTACATGTTCCCGCATGAGTTG 316  
 DB 248 ATGCTACATATTAAGAAAGAA---GAAATTAAGTATATATGCTTCTTGAATTA 304  
 QY 317 TGAATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376  
 DB 305 TCCCTTGAATTTTGAAGCAGTGTAGCACTTACGAAAGTAAAGATTTACTTAGCAGAA 364  
 QY 377 TAAAGTTGAGAAATTAATTAATGATCTATTTGATACGTTTACCGCTTACTGATAT 436  
 DB 365 TCAACATGCGCGATTTTAAATTTTACGCGAAATATGCAAGCTCTCTTCTTACAGCTTA 424  
 QY 437 TGTCAACCGGACCGGCGAAACCTGAAGTAAACCAAGAGCAGCGGCTCAAGTCT 496  
 DB 425 TTGCAAGTAAAGACGATACATTAATGTTGTAAGACAGCAAGATGAATGATATTT 484  
 QY 497 ATGATATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556  
 DB 485 ATGATATCAAGTGTGCTGCTTAACTAATATCAACATTTCCAAAGCAATTAATTAAT 544

Oy	557	TGCACCAATATACAGAAATCAGACCGAGCAATTAGAGACCAAGAGATGGCGGATTAG	61.6
Dd	545	TAAATPACTATCTGTACAGTATCTCCAAAATAATGCTTAAAAATTAATCTTCAGATPAACTAA	60.4
Oy	617	CCCTTTCCTGTCTTTGGCCAAAGCGCTTAGAATGTTGGTCTGCCGGGGGATTTATACACCG	67.6
Dd	605	AATAGGCTGGCTACACCGCTGGATTAGGGTCTCAACACTTACCAGAGTGGAAATGGATTCTG	66.6
Oy	677	CTTCCCGGTTTGTCAAGCGTGTATTATTGAAAGAACATCTTAGAGCCGGCCGATGAAA	73.6
Dd	665	AATCAGCTTTTGTCAAGAGTAGCTTTCAATPAAATTTAATGCTCCAAATGCTGAAAACCGAAG	72.4
Oy	737	CGAAAGGTGTACAGCGCGTTTCAATTTGGCAATATAGAGATPCCAAAGGGCGCAG	79.6
Dd	725	AGGAAATATTTGATTACTTACTTCCACATTTTACATTCGCTGGTGAACAAACAAAGGGACTGG	78.4
Oy	797	TGATTAACGAGAGAGACGAATTCATTATATCGCAATATATCTCCGTGATGTGCAACGAAA	85.6
Dd	785	ATGAAGTTGGTCCAACTCATTTGAATATATCAATTTATTTGTGATGGAACTPACTTACACA	84.4
Oy	857	CTGGAATACTATATTTCACCACTATGACAAATGGCGAAATCCAAAAAGTAAATTATTTTC	91.6
Dd	845	AAGGATATTTTCTACTACCACTTATTCAAACAAACAAATTAACGTTGTGATATGATAA	90.4
Oy	917	ATGAAGACCTTGA	92.9
Dd	905	AGGAAGATCTAGA	91.7

```

RESULT 5
US-11-074-176-369
; Sequence 369, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kleenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altemann, Eric
; APPLICANT: McAniff, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(975)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: ORF 1078; bshb
US-11-074-176-369

```

Query Match	8.4%	Score 83.8	DB 12	Length 975
Best Local Similarity	45.4%	Pred. No. 8.9e-14		
Matches 341; Conservative	0	Mismatches 407	Indels 3	Gaps 1

QY 181 AAACTGGAAATATATTTATTTGCGACGGCATTAAGAAGCCGTTATCTTGTGGCGC 240  
 |||||  
 Db 172 AATGATGATTAATTCATTATTAATGTGATGCAATTATGAAAAGGCGCTGTATTCAGCA 231  
 |||||  
 QY 241 CTTTATTTCCGGCGTATGCGGAGTACGAAAAAGATACGGGAAAGTACCGTTACATT 300  
 |||||  
 Db 232 TTAAATTTTCACAG---TCTATATCAATTAATCTTCTCTAAATCGAAGGTAGAAGAAATAT 288  
 |||||

Oy	30	GTCCCGCAGAGTTTGTGACATGGGTGCTGCAGTCTGCAGTCTTTGGAAAGCGTAAA	348
Db	289	GCTTCTTTGGATTANTGCGANACTTATTAAAGTATGTGAAAATACTGACGATGTAAA	348
Oy	361	GAAGAATTCGATCTTTAACGATGTGAGAAAAAATTAGATCTATTGGATCAGTTTA	420
Db	349	GAATCTTGATTAAGCAAAATATTTTAAATTTAGCTTTTCAGCAAAATTTCTCGCAGCT	408
Oy	421	CCGCTTCACGTGATATTCAGACCCGACCGGCGGAAACCTGACGATAGAACCAAGACA	480
Db	409	GATTTACATTTGATTTTAAAGTATTAAGCTGGTAAAGTATCGTAGTTGATCAACCAT	468
Oy	481	GACGCGCTCAAAAGTCTATGATTAATCCAACTGGTGTCAATGACGAACAGCCCGACTTATA	540
Db	469	TCAGTTTTCATATTTAATGATATATCACTGATATGCTTAACTAACAAATCTGAAATTTCCG	528
Oy	541	TGCGATGTAAACCAATCTGCAGCAATATACAGAAATACGACCGAAGCAATTAGAGACAAA	600
Db	529	GATCAATTAATTAATTAATTAAGTACACGCCGACGTTACTCCACATTAATCTTAAGAAATACA	588
Oy	601	GAAATGGGGGGAATTAGCCCTTTCGCTTTTGGCCAAAGCTTAGGAACGTGTGGCTGGCG	666
Db	589	TTGGTCTCTAATGTGATCTTAAATCTAATATAGTAGAGGCTTAGGATCTACACCAATTACT	648
Oy	661	GGGGATTAATACACCGCTTCCCGGTTTTCAGAGCTGTTATTTTGAAGAACAATTAGAG	720
Db	649	GGTGAATGGAATTTAGCTCTCGATTTTGTTAAGTAGCTTTTGTCTGGACACACTCCA	708
Oy	721	CCGCGCGCCGATGAAGAAAGAAAGGTGTAAACACGCTTTTCAAAATTTTGGCAAAATATGAG	780
Db	709	CAGGAAAAAATGAAGTGAAGAAATGTACTAATTAATTTTCCATATTTCTGCATCTAGTGA	768
Oy	781	ATACCAAGGGCGCAGTGAATACGGAAGAAAGCAAAATTCATTATVACGCAATATCTCC	840
Db	769	CAACCTGATGGTTTGAATGAAGTAAAGAAATATCGCTATGATATTAATATGATATACGAT	828
Oy	841	GTCATGTGCAAGAAACTGGAAACATCAATTTCCACCACTATGACAAATCGGCAATCCAA	900
Db	829	TGATAGACTTAGATTAAGATATTTTGTACTTTACTACTTATATGACATTAATCGGATTAAT	888
Oy	901	AAAGTAATTTATTTCAATGAGACCTTGAAT	931
Db	889	GCAGTAGATATGATTAAGCAAAATTAATTAATTT	919

RESULT 6  
 US-10-750-185-59663/c  
 Sequence 59663, Application US/10750185  
 Publication No. US20050260603A1  
 GENERAL INFORMATION:  
 APPLICANT: MMT GENOMICS, INC.  
 APPLICANT: Denise, Sue K.  
 APPLICANT: KERR, Richard  
 APPLICANT: ROSENFELD, David  
 APPLICANT: HOLM, Tom  
 APPLICANT: BATES, Stephen  
 APPLICANT: PANTIN, Dennis  
 TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
 FILE REFERENCE: NMII100-2  
 CURRENT APPLICATION NUMBER: US/10/750,185  
 CURRENT FILING DATE: 2003-12-31  
 PRIOR APPLICATION NUMBER: US 60/437,482  
 PRIOR FILING DATE: 2002-12-31  
 NUMBER OF SEQ ID NOS: 64922  
 SOFTWARE: PatentIn version 3.1

ORGANISM: Bovine 19866881284323  
US-10-750-185-59863

Query Match 3.9%; Score 38.8; DB 8; Length 720;

Best Local Similarity 57.4%; Pred. No. 0.7;  
Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 824 ATACGCAATTAATCTTCGATGATGCAACGAACTGAACTATTTCCACCACTATG 883  
DB 262 ATACCAATGTTGTCTATCTATTTATATATGCAAGATATATTTATCTATATC 203  
QY 884 ACATCGGCAATCCAAAGTAATTTATTTTATGTAAGACCTTGACTGTTGAGCCTA 943  
DB 202 ACCTGGGTATATAGAAATGCAATTTTATTCAGATTGTTGCTAACTGTTCAACTC 143  
QY 944 AA 945  
DB 142 AA 141

## RESULT 7

US-10-750-623-59863/c  
; Sequence 59863, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-1  
; CURRENT APPLICATION NUMBER: US/10/750,623  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 59863  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Bovine 19866881284323  
US-10-750-623-59863

Query Match 3.9%; Score 38.8; DB 8; Length 720;  
Best Local Similarity 57.4%; Pred. No. 0.7;  
Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 824 ATACGCAATTAATCTTCGATGATGCAACGAACTGAACTATTTCCACCACTATG 883  
DB 262 ATACCAATGTTGTCTATCTATTTATATATGCAAGATATATTTATCTATATC 203  
QY 884 ACATCGGCAATCCAAAGTAATTTATTTTATGTAAGACCTTGACTGTTGAGCCTA 943  
DB 202 ACCTGGGTATATAGAAATGCAATTTTATTCAGATTGTTGCTAACTGTTCAACTC 143  
QY 944 AA 945  
DB 142 AA 141

## RESULT 8

US-11-096-568A-29311  
; Sequence 29311, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 29311  
; LENGTH: 1272

TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(1272)  
; OTHER INFORMATION: Ceres Seq. ID no. 4809357  
US-11-096-568A-29311

Query Match 3.6%; Score 36.2; DB 9; Length 1272;  
Best Local Similarity 57.5%; Pred. No. 5.2;  
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 149 AGACACATGACCGTTTATCGGTATGAGGAGAACTTGAAATATATATTTGCCGACG 208  
DB 419 AGACTGACCTTGGTTCTCGATGATGATGACAAAGCTTGATTCATTCCTTGCCCTT 478  
QY 209 GCATTAATGAAAGCGTTTATCTTGTGCGCGCTTATTTTCCGGCTATGCG 261  
DB 479 GCTTCAATGATGACGATGATCTCTCCGCGCTTATCTATCTTACGTTG 531

## RESULT 9

US-09-925-065A-392869/c  
; Sequence 392869, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 392869  
; LENGTH: 583  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-392869

Query Match 3.6%; Score 35.8; DB 6; Length 583;  
Best Local Similarity 49.7%; Pred. No. 4.6;  
Matches 88; Conservative 1; Mismatches 88; Indels 0; Gaps 0;

QY 802 ACGAAGAAGCAAAATTCATTAATCCGATATATCTTCGATGATGCAACGAACTGCA 861  
DB 341 ATGGAAGAAGCTTAATATATATATATCCAGATTAAGAAAGAAATCTGAAAGCTGAT 282  
QY 862 AACTACTATTTCCACCACTATGACAAATCGCAATTCAGAAAGTAATTTATTTATGATA 921  
DB 281 CTGATATGATTCACCACTATATGACATTCAGAAAAACAAACATGTAATATGTAATAA 222  
QY 922 GACCTGACTGTTGAGAGCTTAAGTGTCTTCGCTAAAGCAAGAGAGATTCAT 978  
DB 221 GATAGTGTTGCGGAGGAGAAAGCGGAGATGAACAGTGCGGACAGAGATTTCTT 165

## RESULT 10

US-09-925-065A-654182  
; Sequence 654182, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

```

; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIORITY FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 654182
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-654182

Query Match
Best Local Similarity 3.6%; Score 35.4; DB 6; Length 419;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```

QY 333 AGCTGTCAGTCTTGGAGACGTAAAGAAAGATTGATCTTAAAGATTGAGAA 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 244 AGATTACAGAAATTTAAGAAACAGAAATTTAAGATATATACGAACTTTATGGA 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 393 AAAATTAGATCTATTGATACAGTTTACCGCTTCACTGATTTGCA 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 TAAATTAGATTGAACAGATTATTTTCTTCCATCTGATTTGTTA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 11
US-09-925-065A-654183
; Sequence 654183, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIORITY FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 654183
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-654183
```

```

Query Match
Best Local Similarity 3.6%; Score 35.4; DB 6; Length 419;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```

QY 393 AAAATTAGATCTATTGGATACAGTTTACCGCTTCACTGATATTTGCA 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 TAAATTAGATTGAACAGATTATTTTCTTCCATCTGATTTGTTA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 12
US-10-932-182A-2969
; Sequence 2969, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; PRIORITY FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2969
; LENGTH: 5901
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-2969

Query Match
Best Local Similarity 3.6%; Score 35.4; DB 7; Length 5901;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
```

```

QY 693 AGCTGTTATTGTAAGACATCTAGACCGCGCCGATGAACGAAGGTGAACAC 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 894 AGCTTATACAAAGAAATGATTTGATTTGGCCCTACAGTGCAGGTAAACAGA 953
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 753 CGCTTTCAATTTTGGCAATATGACGATACCAAGGGCGAGTGAACGAAGAGA 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 954 CGTTCAGTACTTACGATTTAACAACGGTAAACAGTTTCTCGTATCAACGAGAAA 1013
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 813 CGAAATTCATTATACGCAATATCTCCGTGATGTGCAACGAAC 857
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1014 CGAAATTCATTATACGCAATATCTCCGTGATGTGCAACGAAC 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 13
US-10-932-182A-2969
; Sequence 2969, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; PRIORITY FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2969
; LENGTH: 5901
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-2969
```

```

Query Match
Best Local Similarity 3.6%; Score 35.4; DB 7; Length 5901;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
```



QY 753 CGCTTTCAATTTGGCAATATAGCATACCAAGGGCGCAGTGTATACGAAAGCA 812  
DB 954 CGTTGCACTACTTACGATATTAAACAGGTAAACGTTTCCTGATCAACGAGAAA 1013  
QY 813 CGAATTCAATTATACGCAATATATCTCCGTGATGTGCAAGCAAC 857  
DB 1014 CGAATTGATATATCAATATAGCATAGTTTCAAGGTTGTCTACTAGTAC 1058

## RESULT 14

US-11-072-175-102  
; Sequence 102, Application US/11072175  
; Publication No. US20060029944A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT  
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS  
; FILE REFERENCE: D0273A CIP  
; CURRENT APPLICATION NUMBER: US/11/072,175  
; PRIOR FILING DATE: 2005-03-05  
; PRIOR APPLICATION NUMBER: US 60/406,385  
; PRIOR FILING DATE: 2002-08-27  
; PRIOR APPLICATION NUMBER: US 10/648,593  
; PRIOR FILING DATE: 2003-08-26  
; NUMBER OF SEQ ID NOS: 571  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 102  
; LENGTH: 961  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-072-175-102

Query Match 3.5%; Score 35; DB 9; Length 961;  
Best Local Similarity 47.1%; Pred. No. 10;  
Matches 107; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 293 TTCAATGTCTCCGATGATGTTTGACATGGTGCTGTCAGTCTGTGGAAG 352  
DB 466 TTAACACAGGTTCATGAAAGTTGGCTTCTGTGATGTCTGTGGCTGGAACC 525  
QY 353 ACGTAAAGAAAGATTCATCTTAACATGTAAGAAAAAATTAGATCATGGATA 412  
DB 526 AGGAGCAGCACTATGAGATGTTCTGTCTTACGTAAGAAATTTATTTCAATGAGA 585  
QY 413 CAGTTTACCGCTTCACTGATATTTGTGACCCGAGCGGCGGAACCTGACATAGAAC 472  
DB 586 CATTTTCTTATCCACAGTATTTTGTGACACTGTCAATCATGAACCTACCCCTTAGGAA 645  
QY 473 CAAGCAGACGGCTCAAGTCTATGATATCAACCTGGTGTATG 519  
DB 646 AATAAGATTACTGTGCAAAATAAGAGAAACGGAACGTAACGTATG 692

## RESULT 15

US-09-925-065A-392868/c  
; Sequence 392868, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 392868  
; LENGTH: 583  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-392868

Query Match 3.5%; Score 34.6; DB 6; Length 583;  
Best Local Similarity 49.7%; Pred. No. 10;  
Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 802 ACGAAGAAGCAAAATTCATTATACGCAATATATCTCCGTGATGTGCAAGCAAACTGGA 861  
DB 341 ATGGAAGAGCTTAATATGATATTTACAAAGTAAAGACAACTGAAAGACTGCATA 282  
QY 862 AACTACTATTTCCACCACTATGACAAATCGCAAAATCCAAAATTAATTTATTTCAATGAA 921  
DB 281 CTGATATGATTCACACTATATGACATTCAGAAAACAAAACATGTAATATGTAATAA 222  
QY 922 GACCTTGACTGTTTGAGCTTAAAGTGTTCCTGCTAAAGCAAGAGATTTTCAAT 978  
DB 221 GATCAGTGTGTGCGCAGGGGAAACGGGGGRTGAACAGGTGCGCACAGAGATTTCTT 165

Search completed: March 18, 2006, 12:06:10  
Job time : 453 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:44:38 ; Search time 5413 Seconds  
(without alignments)  
10459.279 Million cell updates/sec

Title: US-10-812-387-1

Sequence: 1 catcagtcgacgaagctctac.....atgagcttaattaagatcc 996

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_in: \*  
3: gb\_env: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pac: \*  
7: gb\_ph: \*  
8: gb\_pr: \*  
9: gb\_ro: \*  
10: gb\_sts: \*  
11: gb\_sy: \*  
12: gb\_un: \*  
13: gb\_vl: \*  
14: gb\_hlg: \*  
15: gb\_pl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	996	100.0	996	6	CS128791
2	978.8	98.3	2351	1	D45912
3	978.8	98.3	203901	1	BSUB00021
4	440	44.2	110000	1	AE017333_40
5	440	44.2	110000	1	CP000002_40
6	439.8	44.2	990	6	AX433304
7	293.2	29.4	110000	1	AE017225_35
8	293.2	29.4	110000	1	AE017334_35
9	293.2	29.4	291030	1	AE017036
10	290.6	29.2	110000	1	AE017355_35
11	274.6	27.6	304708	1	AE017001
12	273.8	27.5	287765	1	AE017276
13	233.6	23.5	280050	1	AL591975
14	233.6	23.5	349980	6	AX641665
15	232.6	23.4	990	6	AX413775
16	232.6	23.4	990	6	AX415691
17	196.4	19.7	1087	1	CP020191
18	196.4	19.7	1087	6	CQ874237

19	196.4	19.7	110000	1	BA000016_08	Continuation (9 of
20	177	17.8	10647	1	AE006413	AE006413 Lactococ
21	173.8	17.4	1616	1	BACENVA	M15660 B.sphaeriu
22	167.6	16.8	110000	1	AP008934_01	Continuation (2 of
23	162.2	16.3	110000	1	CP000029_02	Continuation (3 of
24	159	16.0	990	6	AR483788	AR483788 Sequence
25	159	16.0	990	6	AX141691	AX141691 Sequence
26	159	16.0	2975	1	AF269303	AF269303 Staphyloc
27	159	16.0	2975	6	AR485259	AR485259 Sequence
28	159	16.0	2975	6	AX144623	AX144623 Sequence
29	159	16.0	4114	1	AF270144	AF270144 Staphyloc
30	159	16.0	4114	6	AR486098	AR486098 Sequence
31	159	16.0	4114	6	AX145462	AX145462 Sequence
32	157.4	15.8	300242	1	AE016745	AE016745 Staphyloc
33	149	15.0	298050	1	AL935261	AL935261 Lactobaci
34	146	14.7	999	6	AX619460	AX619460 Sequence
35	146	14.7	11466	6	AR354326	AR354326 Sequence
36	146	14.7	11466	6	AR353882	AR353882 Sequence
37	146	14.7	110000	1	BA000017_03	Continuation (4 of
38	146	14.7	110000	1	BA000018_03	Continuation (4 of
39	144.4	14.5	110000	1	CP000046_02	Continuation (3 of
40	144.4	14.5	110000	1	CP000046_03	Continuation (3 of
41	141.2	14.2	110000	1	BA000033_02	Continuation (3 of
42	141.2	14.2	110000	1	BX571856_03	Continuation (4 of
43	141.2	14.2	110000	1	BX571857_02	Continuation (3 of
44	126.4	12.7	343050	1	AL935252	AL935252 Lactobaci
45	120.8	12.1	110000	1	AP006716_02	Continuation (3 of

## ALIGNMENTS

RESULT 1  
LOCUS CS128791 996 bp DNA linear PAT 02/AUG-2005  
DEFINITION Sequence 1 from Patent WO2005066336.  
ACCESSION CS128791  
VERSION CS128791.1 GI:71794812  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1  
AUTHORS Sivaraman, H.  
TITLE Process for production of large amount of penicillin v acylase  
JOURNAL Parent: WO 2005066336-A 1 21-JUL-2005;  
Council of Scientific and Industrial Research (IN); University of  
York (GB)  
FEATURES  
source  
1..996  
/organism="Bacillus subtilis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:1423"

## ORIGIN

Query Match 100.0%; Score 996; DB 6; Length 996;  
Best Local Similarity 100.0%; Pred. No. 1.9e-259;  
Matches 996; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CATATGTCACAGCTTCTTACATTGGAACCTGTCACCGTAAACATGATATTAGCAAGACA 60  
1 CATATGTCACAGCTTCTTACATTGGAACCTGTCACCGTAAACATGATATTAGCAAGACA 60  
DB 1 CATATGTCACAGCTTCTTACATTGGAACCTGTCACCGTAAACATGATATTAGCAAGACA 60  
QY 61 ATGATTTTGATTTTACCTGCGGACAGAGGATTCCTATCCGCGCGCTTACAGCTGG 120  
61 ATGATTTTGATTTTACCTGCGGACAGAGGATTCCTATCCGCGCGCTTACAGCTGG 120  
DB 61 ATGATTTTGATTTTACCTGCGGACAGAGGATTCCTATCCGCGCGCTTACAGCTGG 120  
QY 121 AACAGTGAAGCTGACGGAAGGCCCATCAGACACAGTACGCGTTATCGGTATGCGGAGA 180  
121 AACAGTGAAGCTGACGGAAGGCCCATCAGACACAGTACGCGTTATCGGTATGCGGAGA 180  
DB 121 AACAGTGAAGCTGACGGAAGGCCCATCAGACACAGTACGCGTTATCGGTATGCGGAGA 180  
QY 181 AACCTGGAATATATTTATTTTGGCGACGCGCATTAATGAAGCGGTTATCTTGTGCGGC 240  
181 AACCTGGAATATATTTATTTTGGCGACGCGCATTAATGAAGCGGTTATCTTGTGCGGC 240

Db	181	AAACTGGAAATATATATATTTCGCGACCGGCATTAAATGAAAGCGGTTTATCTTGTCGGGG	240
Oy	241	CTTTATTTTTCGGGGCTATGCGGAGTACGAAAAAAGATACGGGAAGATACCGTTCACAT	300
Db	241	CTTTATTTTTCGGGGCTATGCGGAGTACGAAAAAAGATACGGGAAGATACCGTTCACAT	300
Oy	301	GTCCCGCATGAGTTTGTGACATGGGGCTGTCAGTCTGTCACTCTTGGAAACGTTAAA	360
Db	301	GTCCCGCATGAGTTTGTGACATGGGGCTGTCAGTCTGTCACTCTTGGAAACGTTAAA	360
Oy	361	GAAGAAGTTCGATCTTTAAACGATTGTAGAAAAAATTAGATCTATTGGATACAGTTT	420
Db	361	GAAGAAGTTCGATCTTTAAACGATTGTAGAAAAAATTAGATCTATTGGATACAGTTT	420
Oy	421	CCGCTTCACTGATATTTGTCAACCGCGACGGGCCGAAACCTACGATGAAACCAAGAGA	480
Db	421	CCGCTTCACTGATATTTGTCAACCGCGACGGGCCGAAACCTACGATGAAACCAAGAGA	480
Oy	481	GACGGCCTCAAAAGCTATGATATCAACCTGGTGTCAATGCGAAGACGCCCCGACTTATA	540
Db	481	GACGGCCTCAAAAGCTATGATATCAACCTGGTGTCAATGCGAAGACGCCCCGACTTATA	540
Oy	541	TGCGATGTAAACCAATCTGACGCAATATACAGGAATCAGCCGAGCAATTAGAGCAAA	600
Db	541	TGCGATGTAAACCAATCTGACGCAATATACAGGAATCAGCCGAGCAATTAGAGCAAA	600
Oy	601	GAGATGGGCGGATTAGCCCTTTCTGCTTTTGGCCAAAGCTTAGGAACTGTGTGTCGGC	660
Db	601	GAGATGGGCGGATTAGCCCTTTCTGCTTTTGGCCAAAGCTTAGGAACTGTGTGTCGGC	660
Oy	661	GGGGATTTATACACCCGCTTCCCGGTTTGTCAAGCTGTTATTGTGAAGAATCTAGAG	720
Db	661	GGGGATTTATACACCCGCTTCCCGGTTTGTCAAGCTGTTATTGTGAAGAATCTAGAG	720
Oy	721	CCGGCGGCCGATGAACGAAAGGTGTACAGCCGCTTTCAAAATTTGGCAAAATAGAC	780
Db	721	CCGGCGGCCGATGAACGAAAGGTGTACAGCCGCTTTCAAAATTTGGCAAAATAGAC	780
Oy	781	ATACCAAAAGGGGCGAGTAAACGGAAGAAAGCGAAATCATTTATCGCAATATACCTCC	840
Db	781	ATACCAAAAGGGGCGAGTAAACGGAAGAAAGCGAAATCATTTATCGCAATATACCTCC	840
Oy	841	GTGATGTGCAACGAAACTGGAACCTACTATTTTCCACCACTATGACAAATGGCAAAATCCA	900
Db	841	GTGATGTGCAACGAAACTGGAACCTACTATTTTCCACCACTATGACAAATGGCAAAATCCA	900
Oy	901	AAAGTAAATTTATTTTCATGAAAGCTTGACTGTTTGGAGCCTTAAAGTGTTTTCCGTTAA	960
Db	901	AAAGTAAATTTATTTTCATGAAAGCTTGACTGTTTGGAGCCTTAAAGTGTTTTCCGTTAA	960
Oy	961	GCAGAAAGAGATTTCATGAGCTTATTAATTAAGGATCC	996
Db	961	GCAGAAAGAGATTTCATGAGCTTATTAATTAAGGATCC	996

RESULT 2	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
D45912	D45912	2351 bp. DNA	linear	ECOT 10-FEB-1992				
		Bacillus subtilis genome sequence between the <i>iol</i> and <i>hut</i> operon, partial and complete cds.						
	D45912.1	GI:1408482						
		hutC; pdp; nucC; dra; deoR; yxxB; yxEH; yxEQ; yxEP; yxEO; yxEN; yxEL; yxEL; yxEG; yxEL; yxEL; yxEG; yxEF; yxSE; yxSE; yxEC; yxEL; yxDM; yxDK; hypothetical DNA-binding protein; deoxyribose-phosphate aldolase; pyrimidine nucleoside transport protein; pyrimidine nucleoside phosphorylase; putative histidine						
		disease.						
		Bacillus subtilis						
		Bacillus subtilis						
		Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;						
		1 (sites)						
		Yoshida,K., Sano,H., Miwa,Y., Ogasawara,N. and Fujita,Y.						

TITLE	Cloning and nucleotide sequencing of a 15 kb region of the <i>Bacillus subtilis</i> genome containing the <i>iol</i> operon
JOURNAL	Microbiology 140 (Pt 9), 2289-2298 (1994)
REFERENCE	7952181
AUTHORS	2 (sites)
TITLE	Yoshida, K., Sano, H., Seki, S., Oda, M., Fujimura, M. and Fujita, Y.
JOURNAL	Cloning and sequencing of a 29 kb region of the <i>Bacillus subtilis</i> genome containing the <i>hut</i> and <i>wepA</i> loci
REFERENCE	Microbiology 141 (Pt 2), 337-343 (1995)
AUTHORS	7704263
TITLE	3 (sites)
JOURNAL	Yoshida, K., Fujimura, M., Yanai, N. and Fujita, Y.
REFERENCE	Cloning and sequencing of a 23-kb region of the <i>Bacillus subtilis</i> genome between the <i>iol</i> and <i>hut</i> operons
AUTHORS	DNA Res. 2 (6), 295-301 (1995)
TITLE	8867804
JOURNAL	4 (bases 18222 to 23351)
REFERENCE	Saxild, H. H., Andersen, L. N. and Hammer, K.
AUTHORS	dre-mnpC-pdp operon of <i>Bacillus subtilis</i> : nucleotide sequence, induction by deoxyribonucleosides, and transcriptional regulation
TITLE	by the deer-encoded <i>DeeR</i> repressor protein
JOURNAL	J. Bacteriol. 178 (2), 424-434 (1996)
REFERENCE	8550462
AUTHORS	5 (bases 1 to 23351)
TITLE	Fujita, Y.
JOURNAL	Direct Submission
REFERENCE	Submitted (10-FEB-1995) Yasutaro Fujita, Fukuyama University,
AUTHORS	Biotechnology, 985 Sano, Hishidate-cho, Fukuyama, Hiroshima
TITLE	729-02, Japan (E-mail: yfujita@bft.fukuyama-u.ac.jp,
JOURNAL	Tel: 0849-36-2111 (ex. 4615). Fax: 0849-36-2023)
REFERENCE	Location/Qualifiers
AUTHORS	1. .23351
TITLE	source

gene 1. .657  
/gene="yxdx"

CDS  
1. .657  
/gene="yxdx"  
/note="belongs to the histidine protein kinase family of  
adaptive responses in Bacteria"  
/citation={1}  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA08314.1"  
/db\_xref="GI:1408483"  
/translation="KLDAIRVTYNNQWVHQVTPPLSVINLLIQEEDPEFQIKEKVRQ  
IQEPGELLTSLRLDIFERDPKTEAVSLSLELVSVISYKRFIORYRYPMANVCDDH  
QIYDAKMLKRAIGQVVTNAYKVSAGKSDRELNAVPCDEDEPTVLEVDYGIISODI  
KRVFDPYTGNGRRRFEOSTGIGLHVKEITDKLNTVIDISSFGEISTVAFPLTKK  
"

RBS  
<.788. .>796  
801. .1574  
/gene="yxdL"

gene 801. .1574  
/gene="yxdL"  
/note="homologous to FtsE protein of E. coli, belonging to  
the ATP-binding-cassette (ABC) family"  
/citation={1}  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAA08315.1"  
/db\_xref="GI:1408484"  
/translation="MANMLLEVHINKTYKGVSYOALKOISFIEEGEFTVAMPSPGS  
GNTYLLITISTDRPDSGDLINGENPHRLKRTLAHRRKOLGPFODFNLLDTLT  
GNTMLPTLEKEASVWEEKLHGCIYALGTEINLNGTPEVSGGQORALIAAVIH  
KPSLLADEPTEKSKATKQVMETLOSLENDHYTALMTVTHDPVASYCRVYPIKIK  
VVLFNEIRIGENGRQVFEQIILDLVSLMGANDLSSVFL"

RBS  
<.1537. .>1545  
1549. .3417  
/gene="yxdL"

gene /gene="yxdM"



Db 8854 CCGCTTCACTGATATTTGTCAGACCGGAGCGGCGCCGAAACCTGACGATAGAACCAAGAGCA 8913  
 Qy 441 GACGGCCCTCAAGCTCATATGATATATCAACCTGGTGCATGACGACGACCCGACCTTTATA 540  
 Db 8914 GACGGCCCTCAAGCTCATATGATATATCAACCTGGTGCATGACGACGACCCGACCTTTATA 8973  
 Qy 541 TGGCATGTATACCAATCTGACGACCAATATACGAGATGACGACGACCAATATGAGACAA 600  
 Db 8974 TGGCATGTATACCAATCTGACGACCAATATACGAGATGACGACGACCAATATGAGACAA 9033  
 Qy 601 GAGATGGGGGATTAAGCCCTTTCTGCTTTGGCCAAAGCTTAGAACTGTGCTGCCG 660  
 Db 9034 GAGATGGGGGATTAAGCCCTTTCTGCTTTGGCCAAAGCTTAGAACTGTGCTGCCG 9093  
 Qy 661 GGGGATTAATACACCGCTTCCCGGTTTGCACAGCGCTTTATTTGAAACATCTAGAG 720  
 Db 9094 GGGGATTAATACACCGCTTCCCGGTTTGCACAGCGCTTTATTTGAAACATCTAGAG 9153  
 Qy 721 CCGGGCGGCGATGAAAGCAAGGATGACGCGCTTTCAAAATTTGGCAATATGACG 780  
 Db 9154 CCGGGCGGCGATGAAAGCAAGGATGACGCGCTTTCAAAATTTGGCAATATGACG 9213  
 Qy 781 ATACCAAGGCGCGCATGATTAACGGAAGCAAGAAATTCATTATACGCAATATCTCC 840  
 Db 9214 ATACCAAGGCGCGCATGATTAACGGAAGCAAGAAATTCATTATACGCAATATCTCC 9273  
 Qy 841 GTGATGTGACAGCAAACTGTAATCTATTTCCACCACTATGACAAATGCGCAATCCAA 900  
 Db 9274 GTGATGTGACAGCAAACTGTAATCTATTTCCACCACTATGACAAATGCGCAATCCAA 9333  
 Qy 901 AAAGTAAATTTTATTTTATGACGACCTTGAAGCTTTGGAGCCCTTAAAGTGTTCCTGCTAAA 960  
 Db 9334 AAAGTAAATTTTATTTTATGACGACCTTGAAGCTTTGGAGCCCTTAAAGTGTTCCTGCTAAA 9393  
 Qy 961 GCGAAGAGAGATTTTATGATGCTTAATTA 990  
 Db 9394 GCGAAGAGAGATTTTATGATGCTTAATTA 9423  
 RESULT 3  
 BSUB0021/c 203901 bp DNA linear BCT 18-APR-2005  
 LOCUS Bacillus subtilis complete genome (section 21 of 21): from 4010730  
 DEFINITION to 4214630.  
 ACCESSION Z99124 AL009126  
 VERSION Z99124.2 GI:32468840  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 AUTHORS  
 Bacillus subtilis subsp. subtilis str. 168  
 Bacillus subtilis subsp. subtilis str. 168  
 Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 1 (bases 1 to 203901)  
 Kunze, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,  
 Azevedo, V., Bertolo, M.G., Bessestere, P., Bolotin, A., Borczyk, S.,  
 Bortiss, R., Boudier, L., Brans, A., Braun, M., Brignell, S.C.,  
 Bron, S., Brouillet, S., Brusch, C.V., Caldwell, B., Capuano, V.,  
 Carter, N.M., Choi, S.K., Codani, J.J., Conner, I.F., Cummings, N.J.,  
 Daniel, R.A., Denicot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,  
 Emerson, P.T., Entian, K.D., Errington, J., Fabis, C., Ferrari, E.,  
 Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A.,  
 Galleron, N., Ghim, S.Y., Glaeser, P., Goffeau, A., Golligorsky, E.J.,  
 Grandi, G., Guiseppi, G., Guy, B.J., Haga, K., Halesch, J., Harwood, C.R.,  
 Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hult, M.F.,  
 Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,  
 Klaer, Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P.,  
 Koningsstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,  
 Lardinois, S., Maue, C., Medigue, C., Medina, N., Mellado, R.P.,  
 Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,  
 Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Paro, V., Pohl, T.M.,  
 Portet, D., Portolillo, S., Prescott, A.M., Prescan, E., Pujic, P.,  
 Purnelle, B., Rapoport, G., Ray, M., Reynolds, S., Rieger, M.,  
 Rivoita, C., Rocha, E., Roche, B., Rose, M., Sadate, Y., Sato, T.,  
 Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,

TITLE  
 JOURNAL  
 PUBLISHED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source  
 gene  
 CDS  
 gene  
 CDS  
 Sekowska, A., Seror, S.J., Seror, P., Shin, B.S., Soldo, B.,  
 Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K.,  
 Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpetra, P., Tognoni, A.,  
 Tokuchi, V., Uchiyama, S., Vandenbol, M., Vannier, P., Vaissier, A.,  
 Viari, A., Wandut, R., Wedler, E., Wedler, H., Weitznegger, T.,  
 Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,  
 Yoshida, K., Yoshikawa, H.F., Zumelein, E., Yoshikawa, H. and  
 Danchin, A.  
 The complete genome sequence of the gram-positive bacterium  
 Bacillus subtilis  
 Nature 390 (6657), 249-256 (1997)  
 9384377  
 2 (bases 1 to 203901)  
 Kunze, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.  
 Direct Submission  
 Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,  
 Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724  
 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr  
 adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45  
 68 89 48  
 On Jul 7, 2003 this sequence version replaced gi:2636442.  
 This entry contains data from release R16.1 of the Subtilist  
 database. Further data on gene annotation and detailed information  
 about changes from previous releases can be found at  
 http://genolist.pasteur.fr/Subtilist/.  
 Location/Qualifiers  
 1..203901  
 /organism="Bacillus subtilis subsp. subtilis str. 168"  
 /mol\_type="genomic DNA"  
 /strain="168"  
 /db\_xref="taxon:224308"  
 complement(131..859)  
 /gene="bg1s"  
 /locus\_tag="BSU39070"  
 complement(131..859)  
 /gene="bg1s"  
 /locus\_tag="BSU39070"  
 /EC\_number="3.2.1.73"  
 /function="lichen degradation"  
 /note="alternate gene name: bg1, lics"  
 /codon\_start=1  
 /transl\_table=1  
 /product="endo-beta-1,3-1,4 glucanase"  
 /protein\_id="CAB15943.1"  
 /db\_xref="GI:2636453"  
 /db\_xref="GOA:P04957"  
 /db\_xref="InterPro:IPR000757"  
 /db\_xref="InterPro:IPR008263"  
 /db\_xref="InterPro:IPR008264"  
 /db\_xref="InterPro:IPR008985"  
 /db\_xref="Subtilist:Bg10476"  
 /db\_xref="UniProt/Swiss-Prot:P04957"  
 /translation="MPYKRVLLVTLGFMELFVTAASAQTGSGFPDPFNGVNSG  
 FVQRKDGNSGMEFCTRRANNVSTSLGEMALATSPVYNKPDGSGNSVDTYVGL  
 TVNRKSPANQIVSFTYTPGTGTPDEIDIEFLGKDTTKVQNPNTYTNAGNHEK  
 IYDLPDANAHYTAFAWQPSIKYVDGQKHTATNOIPPTPKIMNLMNGTVD  
 EWLGSYGNVPLAHYDWVRYYTK"  
 complement(1155..1988)  
 /gene="l1ct"  
 /locus\_tag="BSU39080"  
 complement(1155..1988)  
 /gene="l1ct"  
 /locus\_tag="BSU39080"  
 /function="required for substrate-dependent induction and  
 catalolite repression of bg1pH"  
 /codon\_start=1  
 /transl\_table=1  
 /product="transcriptional antiterminator (Bg1G family)"  
 /protein\_id="CAB15944.1"  
 /db\_xref="GI:2636454"  
 /db\_xref="GOA:P39805"  
 /db\_xref="Subtilist:Bg10474"  
 /db\_xref="UniProt/Swiss-Prot:P39805"

```

/translation="MKIAKVINNVISVNEQKELVVMGRLAFQKSSGDVDEARI
EKVFTLDNRDVSSEKTLIDYDPIECMEVSEBIHYALQKGLNDSIYSLTDIN
PAIQRNKGDIKALMETKRLYKDEFAIGKALVMYKNTGVSJPEDEAGFLAH
VNALENMENPDIINTKTMOELISVXKHPKIBENESLHYRPFYTHKKEPAOLNG
THMESQDFLDITVEKXKHAVECTKIQTIEREYKHLTSDLLYTHIERVVKO
A"
gene
complement (2084 . 2764)
/gene="yx1p"
/locus_tag="BSU39090"
complement (2084 . 2764)
/gene="yx1p"
/locus_tag="BSU39090"
/function="unknown"
/notes="similar to unknown proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB15945.1"
/db_xref="GI:2636455"
/db_xref="InterPro:IPR000437"
/db_xref="Subtilist:BG1145"
/db_xref="UniProt/Swiss-Prot:P42307"
/translation="MRRIGLCISLTVLVMSACESEGAOMFADCDOKTYKOTAPD
MSSKKODFOALASDROLAIYSSMIOVSEAFDYGICNPETPVYISMTBEADAXN
LESJKLENKSLKONSEVALLOKTRQDNRSRIIOQLTERAGFFETASMTKLINO
VTPKNACKTRQDLQKKQYQYSAESIKNMSIVKKQKADKASFERHLEALLQKOPG
QVRSSELY"
2971 . 4257
/gene="yx10"
/locus_tag="BSU39100"
2971 . 4257
/gene="yx10"
/locus_tag="BSU39100"
/function="unknown"
/notes="similar to unknown proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB15946.1"
/db_xref="GI:2636456"
/db_xref="GOA:P42306"
/db_xref="InterPro:IPR007114"
/db_xref="Subtilist:BG1144"
/db_xref="UniProt/Swiss-Prot:P42306"
/translation="WMKRFPKQNSVWLYPMANSAYSIVVTAVFPLFYKSAASECV
SAAQSTAYLGYTIAISTFILLAMGLPIGLTADYEGCKKFFGVSSAGVSTMLAFI
PSRHWLLLFYVSAIGFSGANVFYDAFVDPKRNMLVARGGLYIGISTPE
IISIVALLQAEITIPYVSAQSOLSEFIAMWMLPTIMIKHVRXYIKKEPHV
INSEKRLGOTMKIRQYRALFLPLAFAFYIDVGITITMSTYSGSLGGSSLLII
LPTQVYAAPPSIITYGKLABRTGKTMLYGIYIMIVCVAYFMETTLDFMLAMLY
ATISQGIQALSRYFAKLVPRKHANEFEGYNI FGKFSIMGPLLAVTALQTLKGSST
AVPSLIILFVIGVILAFVPEBSTDVSSQOONDLPL"
complement (4276 . 5715)
/gene="dead"
/locus_tag="BSU39110"
complement (4276 . 5715)
/gene="dead"
/locus_tag="BSU39110"
/notes="alternate gene name: yx1n"
/codon_start=1
/transl_table=11
/product="ATP-dependent RNA helicase"
/protein_id="CAB15947.1"
/db_xref="GI:2636457"
/db_xref="GOA:P42305"
/db_xref="InterPro:IPR000629"
/db_xref="InterPro:IPR001410"
/db_xref="InterPro:IPR001650"
/db_xref="InterPro:IPR005580"
/db_xref="Subtilist:BG1143"
/db_xref="UniProt/Swiss-Prot:P42305"
/translation="MSHFKNYQISHDILRLBGLGTBPTKVQOSVIPALREKDLIV
KQGTGKTPASRGILPCELAMNDENKPOLILTPPELAVQVEDYINIGRPRKIAQ
AVFGKSPDKQKALEKOKSHIVVGTGPRVLDDHIEKGTLPDDRISYIVIBADMNKG
FLQVEAILIHLPTERTMLFSATLPDIEKLSROYWQNPHEIIVRAAGLTTNIEHA
```

```

VIQVRENEKESLKDVLMTENPDSCLIFCRTEKHNQLTDELDDLGYPCKIHGMQ
EDRDVWNEPREGERYLVATDVARGIDENISLVNYDLPLEKESYVHTGTGRA
GNKGAISPTAPKPRDLADIEYTGIEPIKTEPSQREVRKRPPLAKINDPRESK
KDXSEBANKIMKLYPNFGKKKKIRAYDFGTITAKIDGVSADIDGITIMDNASVEI
LNGGPHLVKMYKMTVKGKOLKVNKNK"
complement (5797 . 6945)
/gene="yx1m"
/locus_tag="BSU39120"
complement (5797 . 6945)
/gene="yx1m"
/locus_tag="BSU39120"
/function="unknown"
/notes="similar to rhamnogalacturonan acetyltransferase"
/codon_start=1
/transl_table=11
/protein_id="CAB15948.1"
/db_xref="GI:2636458"
/db_xref="GOA:P42304"
/db_xref="InterPro:IPR001087"
/db_xref="InterPro:IPR008994"
/db_xref="Subtilist:BG1142"
/db_xref="UniProt/Swiss-Prot:P42304"

Query Match 98.3%; Score 978.8; DB 1; Length 203901;
Best Local Similarity 99.3%; Pred. No. 8.3e-255;
Matches 983; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATATGTCACAGTCTTAATTGGAACTGCTGACCGGTAAACATGATTATGCAAGACA 60
DB 51819 CAAATGTCACAAAGTCTTAATTGGAACTGCTGACCGGTAAACATGATTATGCAAGACA 51760

QY 61 ATGAAATTTGCATTTCAGCTGGGGAACAGAGTGATTTCTATCCGCGCCGTTACAGCTGG 120
DB 51759 ATGAAATTTGCATTTCAGCTGGGGAACAGAGTGATTTCTATCCGCGCCGTTACAGCTGG 51700

QY 121 AACAGTGAAGCTGACGGAAGGCGCCATCAGACA CAGTACCGCTTATCGGTATGGGAGA 180
DB 51699 AATAGTGAAGCTGACGGAAGGCGCCATCAGACA CAGTACCGCTTATCGGTATGGGAGA 51640

QY 181 AAATTTGGAATATATTTATTTGCGGACGGAATTAAGAAAGCGGTTATCTTGTGGCGG 240
DB 51639 AAATTTGGAATATATTTATTTGCGGACGGAATTAAGAAAGCGGTTATCTTGTGGCGG 51580

QY 241 CTTTATTTTCGGGCTATGCGGAGTGCAGAAACATACGGAAGATACCGTTCACTT 300
DB 51579 CTTTATTTTCGGGCTATGCGGAGTGCAGAAACATACGGAAGATACCGTTCACTT 51520

QY 301 GTCCCGCATGAGTTTGTGACATGGGTGCTGTCACTGTCTGTAAGAGCTTAA 360
DB 51519 GTCCCGCATGAGTTTGTGACATGGGTGCTGTCACTGTCTGTAAGAGCTTAA 51460

QY 361 GAAAAATTCGATCTTAAACGATTGTAGAGAAAAATTTAGTCTATTGGATACAGTTT 420
DB 51459 GAAAAATTCGATCTTAAACGATTGTAGAGAAAAATTTAGTCTATTGGATACAGTTT 51400

QY 421 CGGCTTCATGATATTTGTCAAGCGGAGCGGCGAAACCTGACGATAGAGAACCAAGACA 480
DB 51399 CGGCTTCATGATATTTGTCAAGCGGAGCGGCGAAACCTGACGATAGAGAACCAAGACA 51340

QY 481 GACGGCTCAAAAGCTATGATATCAACTGTGTGATGACGAACGCCCGACCTTTATA 540
DB 51339 GACGGCTCAAAAGCTATGATATCAACTGTGTGATGACGAACGCCCGACCTTTATA 51280

QY 541 TGGCAATGTAACCAATCTGACAGCAATTAACGAATACAGACCAAGCAATTAGAGACAA 600
DB 51279 TGGCAATGTAACCAATCTGACAGCAATTAACGAATACAGACCAAGCAATTAGAGACAA 51220

QY 601 GAGATGGGCGGATTAGCCCTTCTGCTTTGGCGAAGGCTTGAAGACTGTTGGCTGCGG 660
DB 51219 GAGATGGGCGGATTAGCCCTTCTGCTTTGGCGAAGGCTTGAAGACTGTTGGCTGCGG 51160

QY 661 GGGGATTAACACCGGCTTCCGGGTTTGTCAAGAGCTGTTATTGGAAGAACATCTAGAG 720
```







Best Local Similarity 65.3%; Pred. No. 2.9e-108;  
Matches 645; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

```

QY 4 ATGTGCAAGCTTCTTACATTGGAACCTGTGACCTGTAACATGTAATGCAAGAACATG 63
Db 1 ATGTGTACAAAGATACAGTTGACGACTGCGGCGGAGCATTTATGCGGAGACGATG 60
QY 64 GATTTTGCACTTACAGTGGGACGAGAGTGTCTCTATCCGCGCGCTTACAGCTGAGAC 123
Db 61 GATTTTGATTTGAGCTTACCGGAGAGTGTCTGTGCAATCCGCGCTGCTACAAATGAAA 120
QY 124 AGTGAAGCTGACGGAAGGCGCCATCAACACAGTACGCGTTATGCGTATGGGAGAAA 183
Db 121 AGTGAAGCGGACGCGACGACGATGCGCGCGGTATGCGTTATGCGGATGGGAGAAA 180
QY 184 CTGGAATATATATTTATTTGCGGACGCAATATGAAAGGTTTATCTTGTGCGGCGCT 243
Db 181 CTCAAGAAATGCTCTGTTTGGCCGATGCTGTAATGAAAAGATGCTTGTGCGGCGCT 240
QY 244 TATTTTCCGGGCTATGCGGAGTACGAAAAACGATACGCGGAATACGCTTACATGTC 303
Db 241 TATTTTCCGGGCTATGCTGTGTATGTAAGAAAGCAAGCAAAAGCAAGCGCATCTTGG 300
QY 304 CCGCATGAGTTTGTGACATGAGTGTGTGCTGCTGTCTGTCTGTCTTGTGAGACGTAAGA 363
Db 301 CCGCATGAGTTTGTGACATGAGTGTGTGCTGCTGTCTGTCTGTCTTGTGAGACGTAAGA 360
QY 364 AAGATTCGATCTTTAAAGATTTAGAGAAAAATTAATGATCTATGATGATCACTTTTACG 423
Db 361 GCGGCTGTTTCTTAAACATTTGTGAGAGAGTCAAGCTTCTGTGATGCGTGAAGCGCG 420
QY 424 CTTCATGGAATATGTCACAGCGGACGCGCGCAACCTGACATGAAACCAAGAGACAGC 483
Db 421 CTGCAATGCTTATTAACGAGCGGCTCGGCGGAGAGTGTGTTGAACCGACGAGAC 480
QY 484 GGCCTCAAGTCTATGATATCAACCTGTGTCTATGACGACGCGGACCTTTATATAG 543
Db 481 GGCATTTCAAAATTCATGATATCCGCTGCGGTGTGACGACGATGCGGATTTCCCTTGG 540
QY 544 CATGTACCAATCTGACGACATATACAGAAATCAAGCCGAAATTAAGAGCAAGAG 603
Db 541 CATTTGACCAATTTGCGCACTTATCGGCTTTCAGCCGCGGACGTTGCCGCAAAAA 600
QY 604 ATGGGCGGATAGCCCTTTCTGCTTTGGCCAAAGGCTTAGAACTTTGGTCTGCCGCGG 663
Db 601 ATGGGAGGCTGACGCTGTGCGCTTTTGTCAAGAGTTCCGCGCTGCCGATTTGCCAGA 660
QY 664 GATTAATACCGGCTTCCGCGTTTGTCAAGCTGTTTATTTGAAAGACATCTAGGCG 723
Db 661 GATTTTACCGCGCATCCGCTTGTCAAGGCTGCAATTTTGAAGACATGAAGCCT 720
QY 724 GCGGCGGATGAAACGAAAGGTGTACAGCGGCTTTTCAATTTTGGCAATTAAGAGAA 783
Db 721 GTGTCCGATGAAACAGAAAGGCTCTGCGGCTTTCATGCTCTTCAATGAAATAT 780
QY 784 CCAAGAGGCGAGTGAATTAACGGAAGAACGAAATTCATATACGCAATATCTCCGTG 843
Db 781 CTAAGAGGCTTATGATGATGATCAAGGAGGAAGATTAACCCCAATATACGGGCGTA 840
QY 844 ATGTGCAAGCTTCTTACATTGGAACCTGTGACCTGTAACATGTAATGCAAGAACAT 903
Db 841 ATGTGCAAGCTTCTTACATTGGAACCTGTGACCTGTAACATGTAATGCAAGAACAT 900
QY 904 GATTAATTTATTTATGAGACCTTGAAGCTTTGAGAGCTTAAAGTGTTCGCTAAGCA 963
Db 901 GTTTGCTTTTATGAGAGACCTGAGACGAAATTAAGGCGTTTCAGCCAAAGCAG 960
QY 964 GAAGAGAGTATTCATGAGCTTAATTA 990
Db 961 GAAGAGAGCTTATCTCTTCAATCA 987

```

RESULT 7

AE017225.35/c  
WPCOMMENT  
Sequence split into 53 fragments LOCUS AE017225 Accession AE017225

Fragment Name	Begin	End
AE017225_00	1	110000
AE017225_01	100001	210000
AE017225_02	200001	310000
AE017225_03	300001	410000
AE017225_04	400001	510000
AE017225_05	500001	610000
AE017225_06	600001	710000
AE017225_07	700001	810000
AE017225_08	800001	910000
AE017225_09	900001	1010000
AE017225_10	1000001	1110000
AE017225_11	1100001	1210000
AE017225_12	1200001	1310000
AE017225_13	1300001	1410000
AE017225_14	1400001	1510000
AE017225_15	1500001	1610000
AE017225_16	1600001	1710000
AE017225_17	1700001	1810000
AE017225_18	1800001	1910000
AE017225_19	1900001	2010000
AE017225_20	2000001	2110000
AE017225_21	2100001	2210000
AE017225_22	2200001	2310000
AE017225_23	2300001	2410000
AE017225_24	2400001	2510000
AE017225_25	2500001	2610000
AE017225_26	2600001	2710000
AE017225_27	2700001	2810000
AE017225_28	2800001	2910000
AE017225_29	2900001	3010000
AE017225_30	3000001	3110000
AE017225_31	3100001	3210000
AE017225_32	3200001	3310000
AE017225_33	3300001	3410000
AE017225_34	3400001	3510000
AE017225_35	3500001	3610000
AE017225_36	3600001	3710000
AE017225_37	3700001	3810000
AE017225_38	3800001	3910000
AE017225_39	3900001	4010000
AE017225_40	4000001	4110000
AE017225_41	4100001	4210000
AE017225_42	4200001	4310000
AE017225_43	4300001	4410000
AE017225_44	4400001	4510000
AE017225_45	4500001	4610000
AE017225_46	4600001	4710000
AE017225_47	4700001	4810000
AE017225_48	4800001	4910000
AE017225_49	4900001	5010000
AE017225_50	5000001	5110000
AE017225_51	5100001	5210000
AE017225_52	5200001	5228663

Continuation (36 of 53) of AE017225 from base 3500001 (AE017225 Bacillus anthracis str)

Query Match 29.4%; Score 293.2; DB 1; Length 110000;  
Best Local Similarity 55.9%; Pred. No. 1.9e-68;  
Matches 556; Conservative 0; Mismatches 438; Indels 0; Gaps 0;

```

QY 3 TATGTCAAGCTTCTTACATTGGAACCTGTGACCTGTAACATGTAATGCAAGAACAT 62
Db 71198 TATGTCTACTAGTTTGAATTAAGACCAAAAAAGTCAAGATCTTTTGGCAAGACAT 71139
QY 63 GATTTTGCACTTACAGTGGGACGAGAGTGTCTCTATCCGCGCGCTTACAGCTGAGAC 122
Db 71138 GATTTTGCACTTACAGTGGGACGAGAGTGTCTCTATCCGCGCGCTTACAGCTGAGAC 71079
QY 123 CAGTGAAGCTGACGGAAGGCGCCATCAGACAGATGAGGCTTATCGGTATGGGAGAAA 182

```

Db	71078	TAAATTAACGGGTGAATTCATTAATAGAAACATGTAACGGTGGAGATGGGTATTAATCA	71012
Oy	183	ACTTGGAAATATATATATTTGCGCGACGGCATTAAATGAAGCGGTTATCTTGTGCGCGCT	242
Db	71018	TCAGAAAGAGATCATTAATGCGGACGGAGTAAATGAAGCAGATATGACATGTCCAACT	70958
Oy	243	TTAATTTCCGGGCTATGCGGAGTACAGAAAAAAAGATACGGGAAGATACCGTTCAATGTG	302
Db	70958	CTATTTTCAGAGATTCGCTACTTTATAGTCAAAGCATATGATGACACACAAAGATTTGGC	70899
Oy	303	CCCCGATGATTTGTGACATGGGTGCTGTCACTGTCACTCTTTGGAAACGCTAAAGA	362
Db	70898	TCCATTTGATTTTGTAACTTGAGCTGACACAATTCATTTGTCTCAAGATTAAAGA	70839
Oy	363	AAAGATTCGATCTTAAACGATTTGATGAAAAAAATTTGATCTATGATACAGTTTAC	422
Db	70838	ATCTGTAGATAGCATTAACCTTTTGGATATACCATTAACCGGATTTTGAAGCTTACCGCAC	70779
Oy	423	GCTTCACTGATATTTGTTCAGACCGGACGGGCGCAACCTGACATATGAACCAAGACAGA	482
Db	70778	ACTACATTTGATTTTGAAGGATTAATATGGGAGATATGATTTGATCTGGATCCGACATGA	70719
Oy	483	CGGCTCAAAGTCTATGATTAATCAACTGTGTCTATGACGACACAGCCCCGACTTATATG	542
Db	70718	AGGATTTAAATTTGATATGATTAACCACTAGGAGTGATACCAATATGTCGGAGTTTAAATG	70659
Oy	543	GCATTTAACCAATCTGACGACAAATTTACAGGATCAGACCGAACATTTTGAAGCGAAAGA	602
Db	70658	GCATTTTCAAAATTTTAAAGACATATATAGGCTTTAAATCCGACCATTTGCGCGCAACAGA	70599
Oy	603	GATGGCGGATTAAGCCCTTCTGCTTTTGGCCAAAGGCTTAAAGAACTGTGTGCTGCGG	662
Db	70598	GTGAGATTAATTTACCATTAATAGTGTCTTTTGGCCAAAGGCTCGGGCTCAATGGGACTTCCAG	70539
Oy	663	GGATTTATACACGGCTTCCGGGTTTGTACAGGCTTTATTTAAAGAACATCTAGAGCC	722
Db	70538	GGATTTTCACTCCCGCATGAGGTTTGTGCGGGCAGCATATGCGCAACAAACATTCACG	70479
Oy	723	GGCGGCGGATGAACGAAGGTGTAAACGCGCTTTTCAAAATTTTGGCAAAATATGCAAT	782
Db	70478	TATAGATATAGCAAGAGGGAAGTATCAGCCCTTTTCAATCTTATCAAAATTTGAGAGT	70419
Oy	783	ACCAAAAGGCGGCACTGTAAACGAAAGAAAGCAAAATTCATTTATAGCAATATATCTTCGT	842
Db	70418	TCTTAAAGGTGAGATTAATACAGAAAGAGGTGATTAATTAACATATATACAAGCGT	70359
Oy	843	GATGTGCACGAATCTGAAACTGTATTTTCCACACTATGACAAATCGGCAATCCAAA	902
Db	70358	AATGTGTATGGAATCCGGAACTATTTATATCATCTTACGATTTGTAGACAAATTTATGC	70299
Oy	903	AGTAATTTTATTTCATGAAGCCTTGACTGCTTTTGAAGCCTTAAAGTGTTTTCGCTAAAGC	962
Db	70298	TGTTTCATTTATTTTCAATGAATAATTTAGATACAGATGAGATTAAGCTATCCGTTCAAG	70239
Oy	963	AGAGAGAGATTCATGAGCTTAATTAAGATCC	996
Db	70238	GAACCAAAAAATTTTATATGAACTATGATAGCC	70205

Sequence split into 53 fragments				LOCUS	AE017334	Accession	AE017334
Fragment Name				Begin	End		
AE017334_00				1	110000		
AE017334_01				100001	210000		
AE017334_02				200001	310000		
AE017334_03				300001	410000		
AE017334_04				400001	510000		
AE017334_05				500001	610000		
AE017334_06				600001	710000		
AE017334_07				700001	810000		
AE017334_08				800001	910000		

ABO17334_09	900001	1010000
ABO17334_10	1000001	1110000
ABO17334_11	1100001	1210000
ABO17334_12	1200001	1310000
ABO17334_13	1300001	1410000
ABO17334_14	1400001	1510000
ABO17334_15	1500001	1610000
ABO17334_16	1600001	1710000
ABO17334_17	1700001	1810000
ABO17334_18	1800001	1910000
ABO17334_19	1900001	2010000
ABO17334_20	2000001	2110000
ABO17334_21	2100001	2210000
ABO17334_22	2200001	2310000
ABO17334_23	2300001	2410000
ABO17334_24	2400001	2510000
ABO17334_25	2500001	2610000
ABO17334_26	2600001	2710000
ABO17334_27	2700001	2810000
ABO17334_28	2800001	2910000
ABO17334_29	2900001	3010000
ABO17334_30	3000001	3110000
ABO17334_31	3100001	3210000
ABO17334_32	3200001	3310000
ABO17334_33	3300001	3410000
ABO17334_34	3400001	3510000
ABO17334_35	3500001	3610000
ABO17334_36	3600001	3710000
ABO17334_37	3700001	3810000
ABO17334_38	3800001	3910000
ABO17334_39	3900001	4010000
ABO17334_40	4000001	4110000
ABO17334_41	4100001	4210000
ABO17334_42	4200001	4310000
ABO17334_43	4300001	4410000
ABO17334_44	4400001	4510000
ABO17334_45	4500001	4610000
ABO17334_46	4600001	4710000
ABO17334_47	4700001	4810000
ABO17334_48	4800001	4910000
ABO17334_49	4900001	5010000
ABO17334_50	5000001	5110000
ABO17334_51	5100001	5210000
ABO17334_52	5200001	5227419

Continuation (36 of 53) of ABO17334 from base 3500001 (ABO17334 Bacillus anthracis str.

Query Match	29.4%	Score 293.2	DB 1	Length 110000
Best Local Similarity	55.9%	Pred. No. 1.9e-68		
Matches 556	Conservative 0	Mismatch 438	Indels 0	Gaps 0
Qy	3	TATGTGCACAGCTCTTACATTGGAACTCTCTGACCGTAAACATGTAATTAGCAAGAACAT	62	
Db	70624	TATGTGTACTAGTTTACATTTAGACACAAAAACGGTCAGCATCTTTTGGCAAGAACAT	70565	
Qy	63	GGATTTTGCAATTCACGTGGGACAGAGGTGATTTCTATCCGCGCGTTACAGCTGAA	122	
Db	70564	GGACTTCACATTTGATGATGATCAAGAAATAATTAATCATTCCTCGACATTCACAGTGA	705050	
Qy	123	CAATGAAGCTGACGAAAGGCCCTCAGACAGATACGGCTTTATCCGATGAGGGAGAA	182	
Db	70504	TAAATATACCGGTGAATCATTTAATACGAACATGCTACGGTCGGAAGGGATTTATATCA	70445	
Qy	183	ACTTGGAAATATATTTATTTGCGACGCGCATTTATGAAGCGGTTATCTTGTGCGGCGCT	242	
Db	70444	TCAAGGAAGATCATTTATGCGCGACGGAATAATAGAGGTTATGACATGTGCAACACT	70385	
Qy	243	TTATTTTCCGGGCTATGCGAGTACGAAAAAAGATACGGGAAGATACCGTTCACTGT	302	
Db	70384	CTATTTTCCAGGATTCGTACTTATATAGTCAAGCATATGACACACACAAAGAAATTTGGC	70325	
Qy	303	CCCGCATAGTTTGTGACATGGGTGCTGTCACTCTGTCAAGTCTTTTGGGAAGACGTAAAGA	362	
Db	70324	TCATTTGATTTTGTACTGTGAGGTGTGACAAATTCATTTGTGTCAAGAGATTAAAGA	70265	

QY 363 AAAGATTCGATCTTTAAGATTTGAGAGAAAATAGATCTATTGGATACAGTTTACC 422  
 DB 70264 ATCTGTAGATGATGATCTTTTGGATATACATTCACGATTTAGACTTACGCCACC 70205  
 QY 423 GCTTCACGTGATTTGTGACAGCCGAGCGGCCGAAACCTGACATGAGAACCAAGAGCGA 482  
 DB 70204 ACTACATTTGATTTAGCGGATTAATGGGAGATTTGCTTACTGATCCGACAGTGA 70145  
 QY 483 CGGCGCTCAAGCTATGATTAATCAACCTGTGTGATGACGAAACGCCCGACTTTATATG 542  
 DB 70144 AGGATTAATAATTTGATGATTAACCACTAGAGATGACGAAATAGTCGAGTTTAAATG 70085  
 QY 543 GCATGTATCAATCTGACGCAATATACGAAATGACCCGAGCAATATGAGAGCAAGA 602  
 DB 70084 GCATTTACAAAATTTAGACAAATATATAGCCCTTAAATGCAAGCCATTCGCCAACA 70025  
 QY 603 GATGGCGGATTAAGCCCTTTCTGCTTTGGCCAAAGGCTTAAGAACTGTGTGCTGCCGG 662  
 DB 70024 GTGAGATTAATTTACATTAATGCTTTGGCCAAAGGCTCGGGCTCAATGGGACTTCCAG 69965  
 QY 663 GGATTTATACACCGCTTCCCGGTTTGCAGAGCTGTTTATTTGAAAGCACTTACAGCC 722  
 DB 69964 GGATTTACACCGCCATGAGGTTTGTGCGGGCAGCATATGCAAAACAAACATTCAGG 69905  
 QY 723 GCGCGCCGATGAACGAAAGGTGTACAGCCGCTTTTCAATTTTGGCAATATGAGCAT 782  
 DB 69904 TATAGATGACGAAAGGAGGATATGACCCCTTTTCAATCTTATCAAAATGTGAGGT 69945  
 QY 783 ACCAAAGGCGCAGATGATACGAGAAAGCAAAATTCATTAACGCAATATCTCCGT 842  
 DB 69844 TCTTAAGGTGAGATTAATACGAAAGGTCATTAATTAATACCAATATACAGAGGT 69785  
 QY 843 GATGTCAACGAACTGAAACTACTATTTCCACCATATGCAATTCGCAATTCGCAAA 902  
 DB 69784 AATGTATGAGATCCGAAACATATTTATCATCTTACGATGATGACAAATATATAGC 69725  
 QY 903 AGTAAATTTATTTCTGAAAGCTGACTGTTTGAAGCTTAAGCTTTCGCCGTAAGC 962  
 DB 69724 TGTTCATTTATTTCTGAAATTTAGATACAGATATTAAGCTTATTCGTTCCAAAG 69665  
 QY 963 AGAAGAGATATTCATGAGCTTAATTAATGATCC 996  
 DB 69664 GAAACAAAATATTTTATGAGAACTAATGATCC 69631

RESULT 9  
 AEO17036/c 291030 bp DNA linear BCT 01-MAY-2003  
 LOCUS AEO17036  
 DEFINITION Bacillus anthracis str. Ames section 13 of 18 of the complete genome.  
 ACCESSION AEO17036  
 VERSION AEO17036.1 GI:30258352  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.  
 AUTHOR Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Oksstad, O., Heigaasen, E., Riststone, J., Wu, M., Kolonay, J., Beaman, M., Dodson, R., Brinkac, L., Gwinn, M., DeBoy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Kouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plant, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.  
 TITLE The genome sequence of Bacillus anthracis Ames and comparison to closely related bacteria  
 JOURNAL Nature 423 (6935), 81-86 (2003)  
 PUBMED 12721629

REFERENCE 2 (bases 1 to 291030)  
 AUTHORS Read, T., Peterson, S., Tourasse, N., Baillie, L., Paulsen, I., Nelson, K., Tettelin, H., Fouts, D., Eisen, J., Gill, S., Holtzapple, E., Oksstad, O., Heigaasen, E., Riststone, J., Wu, M., Kolonay, J., Beaman, M., Dodson, R., Brinkac, L., Gwinn, M., DeBoy, R., Madupu, R., Daugherty, S., Durkin, A., Haft, D., Nelson, W., Peterson, J., Pop, M., Kouri, H., Radune, D., Benton, J., Mahamoud, Y., Jiang, L., Hance, I., Weidman, J., Berry, K., Plant, R., Wolf, A., Watkins, K., Nierman, W., Hazen, A., Cline, R., Redmond, C., Thwaite, J., White, O., Salzberg, S., Thomason, B., Friedlander, A., Koehler, T., Hanna, P., Kolsto, A.-B. and Fraser, C.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
 FEATURES  
 source  
 complement (236808..280850)  
 /organism="Bacillus anthracis phage lambda Ba03"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:229345"  
 /note="Putative prophage; Putative lysogenic prophage region. Best hits not completed. Putative attL/R"  
 complement (119..1390)  
 /locus\_tag="BA3835"  
 complement (119..1390)  
 /locus\_tag="BA3835"  
 /locus\_tag="BA3835"  
 /note="similar to GP:2827439, and SP:P04028; identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="conserved hypothetical protein"  
 /protein\_id="AAP27572.1"  
 /db\_xref="GI:30258353"  
 /translation="MPDRLKNGEKIAPIVKESQITVEHRADEVIESNOFVLESEF RKHISDSHPITPTGVYDIDGRDLEKVAADVAGAGVAPQIISGTHAISVALFG IIRGDBLITITGKPYTLEIEIVYRKNGKISPRENIGVANVLTEGVIDLAVAA AIHSNTMIGIORSKGYATPSPFTISQIKEMAFKIKPDDVVFVFNVCYEFIEBOE PCHVGADLMAGSLIKNPGGIVKNGGIVKEQIVKACARLVSPGIGAEAGSLYL OEMQGFPLAHPVAGALKGAIFPAALFEKIGNTSPAMNAPRDLQSVFDPDKRM IAPCOAIOVASPINSHFTPYANWPGEDVIVMAGFICGASIELSADGPIRBPYVA YVGGELTSHVKIACSAIDALIEKELLTIS"  
 complement (1383..1595)  
 /locus\_tag="BA3836"  
 complement (1383..1595)  
 /locus\_tag="BA3836"  
 /note="identified by Glimmer2; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="hypothetical protein"  
 /protein\_id="AAP27573.1"  
 /db\_xref="GI:30258354"  
 /translation="MKEENRYYOIEIPSESKLLTLKTEFTLLTMEFLDQFYVDCG YIRHSSLNVOLEKPIVERGENNV"  
 complement (1592..2578)  
 /locus\_tag="BA3837"  
 complement (1592..2578)  
 /locus\_tag="BA3837"  
 /note="similar to SP:P04028; identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="GTP-binding protein"  
 /protein\_id="AAP27574.1"  
 /db\_xref="GI:30258355"  
 /translation="MKELASLAKTARABELVSTQKPKFPATYIGKXLEELTMLT EELPDIYVFNELTSPQIRLTSVILARVVDLPQILIDIPOAKSREGLOVSLAO LQYTPRMGGSLSRGGIGTRGPEFTLETRDRIRIRIRIDIKQALVAVVQHRK RYRRRRDNKVFQVSLIGYTNAGSTLENRLTEADTEBENILFATLDTTKMPLPSG





QY	783	ACCAAGGCGCAGATACGGAAGACGAAATTCATTATACGCAATATACCTCGGT	842
Db	95798	TCCTAAAGCGGAGTATATACGAAAGAGTACTTATATATATATATATACAGCCT	95739
QY	843	GATTCGACGAAACTGGAAGTACTATTTCCACGATATGACATCGGCAATCCAAA	902
Db	95738	AATGTGTATGAGTCGCGGAACATCATATATATATATATATATATATATATATG	95679
QY	903	AGTAATTTATTTATGAGACCTTGACGCTTTGGAGCTTAAGTCTTTCCGTAAGC	962
Db	95678	TATTCATTTATTTATGAGAAATTTATGATATACAGATGATTAAGCCTATCCGTCACG	95619
QY	963	AGAGAGAGTATTCATGAGCTTAAATTAAGATC	995
Db	95618	GAAACAAAAATTTTATGAAATTAAGTACG	95586
RESULT 11			
LOCUS	AE017001	304708 bp	DNA linear BCT 23-MAY-2003
DEFINITION	Bacillus cereus ATCC 14579 section 4 of 18 of the complete genome.		
ACCESSION	AE017001	AE016877	
VERSION	AE017001.1	GI:29894619	
KEYWORDS			
SOURCE			
ORGANISM	Bacillus cereus ATCC 14579		
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.		
AUTHORS	1 (bases 1 to 304708) Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapatchal,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goldsman,E., Larsen,N., D'Souza,M., Walunas,T., Greckhin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.		
TITLE	Genome sequence of Bacillus cereus and comparative analysis with Bacillus anthracis		
JOURNAL	Nature 423 (6935), 87-91 (2003)		
PUBMED	12721630		
REFERENCE	2 (bases 1 to 304708) Candelson,B., Galloux,K., Ehrlich,D.S. and Sorokin,A.		
AUTHORS	The number of ribosomal RNA operons in Bacillus cereus		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 304708) Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelson,B., Kapatchal,V., Bhattacharya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goldsman,E., Larsen,N., D'Souza,M., Walunas,T., Greckhin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.		
AUTHORS	Direct Submission		
TITLE	Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Villvert, Jouy en Josas 78352, France		
JOURNAL	Location/Qualifiers		
FEATURES	1..304708 /organism="Bacillus cereus ATCC 14579" /mol_type="genomic DNA" /strain="ATCC 14579" /db_xref="ATCC:14579" /db_xref="taxon:226900"		
gene	85..345 /locus_tag="BC0922"		
CDS	85..345 /locus_tag="BC0922" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP07909.1" /db_xref="GI:29894620"		
gene	433..543 /locus_tag="BC0923"		
CDS	433..543 /locus_tag="BC0923"		
gene	1046..1141 /locus_tag="BC0924"		
CDS	1046..1141 /locus_tag="BC0924" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP07911.1" /db_xref="GI:29894621"		
gene	1415..1582 /locus_tag="BC0925"		
CDS	1415..1582 /locus_tag="BC0925" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical Membrane Spanning Protein" /protein_id="AAP07912.1" /db_xref="GI:29894623"		
gene	1678..1809 /locus_tag="BC0926"		
CDS	1678..1809 /locus_tag="BC0926" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP07913.1" /db_xref="GI:29894624"		
gene	1824..2279 /locus_tag="BC0927"		
CDS	1824..2279 /locus_tag="BC0927" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP07914.1" /db_xref="GI:29894625"		
gene	2515..2712 /locus_tag="BC0928"		
CDS	2515..2712 /locus_tag="BC0928" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP07915.1" /db_xref="GI:29894626"		
gene	2724..3116 /locus_tag="BC0929"		
CDS	2724..3116 /locus_tag="BC0929" /locus_start=1 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAP07916.1" /db_xref="GI:29894627"		
gene	3357..3575 /locus_tag="BC0923"		

```

CDS      /locus_tag="BC0930"
3357..3357
/locus_tag="BC0930"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP07917.1"
/db_xref="GI:29894628"
/translation="MRRKRDNRMOKFSKIVICVGLIMIIIGMIVHDTWKDRIFESLA
VLPTVLSLIMNIYAHNIMDKFRKENK"
3624..4130
/locus_tag="BC0931"
3624..4130
/locus_tag="BC0931"
/codon_start=1
/transl_table=11
/product="hypothetical Membrane Spanning Protein"
/protein_id="AAP07918.1"
/db_xref="GI:29894629"
/translation="MKSLDFNKKAFEYMYALYSVNDIRSTIIITLVIGIADIYVLLPA
FANPQPIYIIVIPVAFINMYAIIIVINPKROLYTFPGVFGIVCSGLAVITG
KPAYMGLQTPRIVIFSVGLYIFALNYYNKHIIKLGPIKKSXSKKFWMSIGF
CRIGSTHC"
4410..4601
/locus_tag="BC0932"
4410..4601
/locus_tag="BC0932"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP07919.1"
/db_xref="GI:29894630"
/translation="MQAFVGLVTFVAIIILVVALSPVNLFKYIIFIGVLIIGA
ISLGANKPELHSHKFNKK"
4661..5077
/locus_tag="BC0933"
4661..5077
/locus_tag="BC0933"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP07920.1"
/db_xref="GI:29894631"
/translation="MVFRCXYILVEGFTIQNRLEHGKQAGVEPPYSNAAHLLTP
WNDKRAIEQELKEFGIHDHSAANGVFLPYKNEVTEVLAHGHCTDYMKEVTKY
LKEVYEGGTQDAVALHDRIIRLLDGLKXNSPK"
complement(5688..6143)
/locus_tag="BC0934"
complement(5688..6143)
/locus_tag="BC0934"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP07921.1"
/db_xref="GI:29894632"
/translation="MSQMLSDNKADHDGCGILKFLGMSIIFILIVPYLMEFQES
NEKPLVNTSSNOANOIEITGSIYGLSDPKYIKMFKSKMLVEKTKRVANFNIN
SSDLKLEISKNTKVIYVMKFEYETKLEVDPEITGIGYINSTNNLL"
complement(6723..8045)
/locus_tag="BC0935"
complement(6723..8045)
/locus_tag="BC0935"
/codon_start=1
/transl_table=11
/product="Citrate-proton symporter"
/protein_id="AAP07922.1"
/db_xref="GI:29894633"
/translation="MAQVKSNNVAGNI FRGSVGNLIEMWDVYSAFVYFSAEFPFK
GPTSQLNTAAIFAVGFLMRPIGSLMRVADRHGRRAALTLISIVAGSLLIAC
TBSYESIGMAPIIVILVARLLOGLSGEGTSATYVSEMASSGRGFYSFOYVTLVA
GOMVANGVOIVIOQLSEPDMAKGRIRPIITGAMGAVAVIMLRMTDSSEDPNSIK
QKESAGTIRALKRHKAVLVYIGLGGTVAFYTTTTLQKMTVTVGLPREVYSMI
NEVALLI FVVLQPIAGLSDKIGRRPLMAFGILGLTLAPIFFFEKTEBETVAFL

```

```

Query Match      27.6%; Score 274.6; DB 1; Length 304708;
Best Local Similarity 54.8%; Pred. No. 2.1e-63;
Matches 544; Conservative 0; Mismatches 449; Indels 0; Gaps 0;

QY      4 ATGTCACAACTCTTACATTGGAACCTGTCACCGTTAAACATGATTAAGCAAGAACATG 63
Db      85486 ATGTCATCTAGTTGAGGTTACAACTAAATTTCCAACTCTTTTGCAGAAACAATG 85545
QY      64 GATTTTGATTTGAGTGGGACAGAGTGATTTCTATCCGCCCGCTTACAGCTGGAAC 123
Db      85546 GATTTACATTAGATTAATGATACAGGAAGTATATCATCTCCGCCGTTACAGTGAAC 85605
QY      124 AGTGAAGCTGACGGAAGGCCCATCAGACACAGTACCGCTTTATCGGTATGGGAGAAA 183
Db      85606 AATATTAAGTGTGAACCAATCCGAACAAAAGGCTGTTCGGAATGGGAATTAATTTT 85665
QY      184 CTGGAAATATATATTTTCCGAGCGCATTAATGAAGCGGTTATCTTGTGCGCGCTT 243
Db      85666 GGAAGAAAGGTCATGATGCGAGATGAGTAATATGAAGCTGTATGACATGTGCAACTTC 85725
QY      244 TATTTTCGGGCTATGCGAGATACGAAAAACGATACGGGAAGATACCGTTCACATTGTC 303
Db      85726 TATTTCCAGATTCGCTACTTATAGTAGTATGATGATGATACCAATTAACGAATGAC 85785
QY      304 CCGATGATGTTTGACATGGGTGCTGTCACTGTCTGCACTTTGGAAAGAGTAAAGAA 363
Db      85786 CCATTTGATTTTGATGACTTGAGCGCTCAGCGATTCATCTGTTGAAATTAAGAAA 85845
QY      364 AAGATTCATCTTTAAGCATTTGTAGAGAAAATTAAGATCAATGATTAAGATTACG 423
Db      85846 TCTATGATGACATTTGCTTAATGATGATGATGATGATGATGATGATGATGATG 85905
QY      424 CTTCAGTATGATGTCAGACCGGACCGGAAACCTGACGATGAAACGAAGCAGAC 483
Db      85906 CTACATTTGATTTTGAAGATTAATTCGGAAGATGATGATGATGATGATGATGATG 85965
QY      484 GGCCTCAAGTCTATGATTAATCAACCTGTGTCTATGACGAACAGCCCGCATTTATG 543
Db      85966 GGAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 86025
QY      544 CATGTAACCATTCGACCAATTAACGAATTCAGACCGAAGCATTTAGAGCAAGAG 603
Db      86026 CATTTACAAATTTAAGCAATTAATTTGCTTAATGCGACCATTTGCGCAACGAG 86085
QY      604 ATGGCGGATTAAGCCCTTCTGCTTTGGCCAAAGGCTTAGAACTGTGTGCGGG 663
Db      86086 TGGGCGCATGATCAATTAAGTCTTTTGGCCAAAGGTCAGGATCGATGGGACTTCCTG 86145
QY      664 GATTATACACCGCTTCCCGGTTGTCAAGCTGTTATTGGAAGAAACATCTAGAGCG 723
Db      86146 GATTTACACCGCCATCAAGATTTGTACGGGAGGATGCGCAACAAACATTCAGGT 86205
QY      724 GCGGCGCATGAAACGAAGGTGTACAGCCGCTTTCAATTTTGGCAATATAGCAT 783
Db      86206 ATAGGAACGAAGAGAGGTATATCGCTATTTTCAATTTTATCGAATGTGAGTT 86265
QY      784 CCAAGGCGCGAGTATACGGAAGACGAATTCATTTATACCAATATATATCTCCGTG 843
Db      86266 CCTAAGGGGCGAGTAAACGAAGACGATATATAGTAATATAGATCATATCAAGTGA 86325
QY      844 ATGTCAACGAACCTGGAACCTACTATTTCCACCACTATGACATGGGCAATCCAAA 903
Db      86326 ATGTATAGAGTCTGGAACATATATTAATCACTATCGATTTGAGCAAAATTAATG 86385
QY      904 GTAATTTATTTCAAGAGACCTTGACTGTTTGAAGCCTTAAGTGTTCCTTAAGCA 963
Db      86386 ATTCATTTATCAATGAAGATTTAATACAGCTGATTAATAACGATTCCTTCAGCG 86445
QY      964 GAAGAGATTTATGATGATTAATTAAGATCC 996
Db      86446 AAACAAAAATTTTATGAAAAACGAAGCGC 86478

```

RESULT 12	AE017276/c	287765 bp	DNA	linear	BCT 23-FEB-2004
LOCUS	AE017276	287765 bp	DNA	linear	BCT 23-FEB-2004
DEFINITION	Bacillus cereus ATCC 10987, section 13 of 18 of the complete genome.				
ACCESSION	AE017276	AE017194			
VERSION	AE017276.1	GI:42738687			
KEYWORDS					
SOURCE	Bacillus cereus ATCC 10987				
ORGANISM	Bacillus cereus ATCC 10987				
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.				
AUTHORS	1 (bases 1 to 287765) Rasko, D.A., Ravel, J., Okerad, O.A., Helgason, E., Cer, R.Z., Jiang, L., Shores, K.A., Fouts, D.E., Tourasse, N.J., Angiuoli, S.V., Kolonay, J., Nelson, W.C., Kolsto, A.-B., Fraser, C.M. and Read, T.D.				
TITLE	The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic adaptations and a large plasmid related to Bacillus anthracis pXOI				
JOURNAL	Nucleic Acids Res. 32 (3), 977-988 (2004)				
PUBMED	14960714				
REFERENCE	2 (bases 1 to 287765) Rasko, D.A., Ravel, J., Okerad, O.A., Helgason, E., Cer, R.Z., Jiang, L., Shores, K.A., Fouts, D.E., Tourasse, N.J., Angiuoli, S.V., Kolonay, J., Nelson, W.C., Kolsto, A.-B., Fraser, C.M. and Read, T.D.				
AUTHORS	Submitted (19-FEB-2004) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
TITLE	Location/Qualifiers				
JOURNAL	1. 287765				
FEATURES	/organism="Bacillus cereus ATCC 10987"				
Source	/mol_type="genomic DNA"				
	/strain="ATCC 10987"				
	/db_xref="ATCC:10987"				
	/db_xref="taxon:222523"				
	complement (138..704)				
gene	/locus_tag="BCE3711"				
CDS	complement (138..704)				
	/locus_tag="BCE3711"				
	/note="identified by Glimmer2; putative"				
	/codon_start=1				
	/transl_table=11				
	/product="hypothetical protein"				
	/protein_id="AAS42616.1"				
	/db_xref="GI:42738688"				
	/translation="MGFLSVILRGILLYVYICLYILWYLAFLPIIIMAHSTLSVTSIPI IMPLVILWIKSITLYAVKRNREKKQMSGVILVILPIVCTAQLSMHTTYSKTKR DRVLHABDKRVHNVDDLQKTKLTKSNEITQLGPIHTRNGENVITTSYLTGER GFLPDSHVLVLFDRDGVKVEYKVNHD"				
gene	complement (711..1730)				
CDS	/locus_tag="BCE3712"				
	complement (711..1730)				
	/locus_tag="BCE3712"				
	/codon_start=1				
	/transl_table=11				
	/product="conserved hypothetical protein"				
	/protein_id="AAS42617.1"				
	/db_xref="GI:42738689"				
	/translation="MDVYILGVILLLLAIVFVLYKPKOLESESQAEPKKAITY ENRIATVTEHKEQQTLLKMEQKTYNDLISAAELENRMKNOLAQVSHKERSEM QEHSHNEIHFOKLADLREYKNGAIVSTHETLQYWKGFVQGIILNEPIMINAV FKNQGRNDRIHHLVLSKTYVLETKEWTKLHGLTKENASTYSPIIDKIGKYQ OEIEKEETEYITGDSLTIOVNGENPVPYRAKLSHILYNCKEMQVDIVKENINI VYFVNGKXEVLDLSEKTPRLKRDQIVSFNNEFLTGKVIYTGQELKRETIINRM NYIMN"				
gene	complement (1871..2917)				
CDS	/locus_tag="BCE3713"				
	complement (1871..2917)				
	/locus_tag="BCE3713"				
	/note="identified by match to protein family HMM PF01183"				
	/codon_start=1				
	/transl_table=11				
	/product="glycosyl hydrolase, family 25, putative"				
	/protein_id="AAS42618.1"				
	/db_xref="GI:42738690"				
	/translation="MODRSSSTNTPFIDVSHMEGNIIMNAVKGSGISAAYAKTEGVNY IDPTFKNVQAAANAVLIGAHFARPEQNDALISEAKYPSIIQSHTDMLPVDLES PTDPSNLTGSAISNMARSPINYQATKEDVWLITGVYINIEFGISGSDIPILMLA KYSTPDPADAGTWTAMQYDTSQGISGVGNDVSAVLEALQNGAGAGGVNSFP NNATPYGVVITINGDNVNLRSGPSVIRQLNREAYEAVWEGOGMLCLGTQWY NDSYIOYKHVYVATITGDVNLRLDAPSLNNGVIRQLHHGSEYRWVSKODGMLCLGTNQ WYVDSYIOTGVQ"				
gene	complement (3049..4812)				
CDS	/locus_tag="BCE3714"				
	complement (3049..4812)				
	/locus_tag="BCE3714"				
	/note="identified by match to protein family HMM PF00005; match to protein family HMM PF00664"				
	/codon_start=1				
	/transl_table=11				
	/product="ABC transporter, ATP-binding/permease protein"				
	/protein_id="AAS42619.1"				
	/db_xref="GI:42738691"				
	/translation="MAEKQSDLRLLSLSTYRKYKGLALFLFVGATVTEMGPPIL KQFDHILVRNPDQALVTLFVYIVAHILKLPYTDLLITQNLAKFVQMRVY YEHVQLSISFPRITGTLVSRITNDTEALIKQFVSIVSTFKNTVFLAGILVAFEL LVNKLAFSLVILPIMFALMVLYRRSAFYLEVRNQLSTLNKAKNESIQGMIVQVF ROEKRRKPEPVNNKHVSAGRTLLDALLRPAPDLVHIVAILVLGLGIDALKS PVEGVLYAVVNYIHRFPQVNMAMKLSPFOALVSSSRVFLMDEKDLAPVQKDG NPQVINGDIEFKNVTFSTYQKDVILKNSPHTVQGVTAFFVGHGSGKSTIMLLRF NINSGNIVIDGELKEPEQERIKKIGVLDQPLPAGVKNINMYEIBLDEBK BAQFOANTFIEKLEPQYETEVEVGAAPSSGQROLINPARTATNPVLDENAT NIDTETEDALQTLQQRKRTTIALHRLSTIQDADQIFVMDHGEIVETGHOELLS EGLVYNYMLIQNKSGLOKAL"				
gene	complement (4812..5563)				
CDS	/locus_tag="BCE3715"				
	complement (4812..5563)				
	/locus_tag="BCE3715"				
	/note="identified by match to protein family HMM PF00005; match to protein family HMM PF00664"				
	/codon_start=1				
	/transl_table=11				
	/product="ABC transporter, ATP-binding/permease protein"				
	/protein_id="AAS42620.1"				
	/db_xref="GI:42738692"				
	/translation="MKVFNNLWMPFKQKRAVITGILLFGVALLVLPKVLGIVVD EINDGTLTDKLLKVVLLVVGITWYILRYWRINIPGGSLGLARLNLYEHFK MSPSFYQNRGTGDMAHNTDIOALTQAGVLTIVDSLAVGCVLVMGPFISML TLLSLIPMIVAIISTNYGTILHKKHKAQGSFSEINDVQESMSMKYIRISQGEK DLOAFKSEEDVHKKMLVARIDSLDFPTIALVGSFLIACVGSVLVVRGELVGD LVTFYVLTGLVWPMILFGLMFINMERGASVDYRKILISQSDVNRNVAVHTIASG DISFAVDSFSYKKNBLHLTDHFDLKKGETLGIYRTGAGTTLKLCILREYDFNG ELKVGSRDIRDVTLYGVRSALISVPQDHLFASIGENIAFGKADLTCEIETAAEIA CIHNDILOFSEGEYTVGERSVLSGGQKORISIRALLTNEIILLDCLSAVDKAT ERTIINALRERAGTGTITTAHRLSAIQANLILVDESRIVQGRTHQMLMENGWYK EMYESQLEALYVEKGV"				
gene	complement (6687..6908)				
CDS	/locus_tag="BCE3716"				
	complement (6687..6908)				
	/locus_tag="BCE3716"				
	/note="identified by match to protein family HMM PF03672"				
	/codon_start=1				
	/transl_table=11				
	/product="conserved hypothetical protein UPF0154"				
	/protein_id="AAS42621.1"				
	/db_xref="GI:42738693"				
	/translation="MPVILGILVAVAGVALGFFIARKTMMNYVLQKNPPINEQML KMMNQMGKQPSQKXINQMSAMKQAK"				
gene	complement (6987..7427)				
CDS	/locus_tag="BCE3717"				
	complement (6987..7427)				
	/locus_tag="BCE3717"				
	/codon_start=1				
	/transl_table=11				
	/product="conserved hypothetical protein"				
	/protein_id="AAS42622.1"				





Dussange, O., Entian, K.D., Fshih, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Haef, J., Jackson, D., Jones, L.M., Kaerst, U., Kieft, J., Kuhn, T., Kunst, F., Kurapkat, G., Madueno, E., Maitouram, A., Vicente, J.M., Ng, E., Nedjari, H., Nordisk, G., Novella, S., de Pablo, B., Perez-Diaz, J.C., Purcell, R., Remmel, B., Rose, M., Schueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J.A., Voss, H., Wehland, J., and Cosset, P.

Comparative genomics of *Listeria species*  
Science 294 (5543), 849-852 (2001)

2 (bases 1 to 280050)  
Glaser, P., Frangeul, L. and Rusniok, C.  
Direct Submission  
Submitted (06-JUN-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE  
E-mail: pglaser@pasteur.fr  
Phone: +33 1 45 68 89 96, Fax: +33 (0) 1 45 68 87 86.  
Location/Qualifiers  
1. 280050  
/organism="Listeria monocytogenes"  
/mol\_type="genomic DNA"  
/strain="EGD-e"  
/db\_xref="taxon:1639"  
158. 584  
/gene="lmo0332"  
158. 163  
/gene="lmo0332"  
172. 507  
/gene="lmo0332"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC98411.1"  
/db\_xref="GI:16409710"  
/transl\_table="UniProt/TrEMBL:O8YA33"  
/translation="MITYTCALTYFIALVSFGESLDALIKSRKNGDALINAKYAVS RSLSLVALGLFIPKSDAPLVALSLVIGALPFGIIGIKISTFKTVGPLTAVGV IMILPLTI"  
565. 584  
/gene="lmo0332"  
936. 6272  
/gene="lmo0333"  
936. 6272  
/gene="lmo0333"  
/note="similar to internalin proteins, putative peptidoglycan bound protein (LPXng motif)"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC98412.1"  
/db\_xref="GI:16409711"  
/db\_xref="GOA:O8YA32"  
/db\_xref="InterPro:IPR001611"  
/db\_xref="InterPro:IPR001889"  
/db\_xref="InterPro:IPR007092"  
/db\_xref="InterPro:IPR009459"  
/db\_xref="UniProt/TrEMBL:O8YA32"  
/translation="MKKFSIVLISVLLGLYALPFDLTLAGDETVSEDTAVKTBAS DAKTEGIESRSDDETAERPKAEKAEKETEPEKAKTEBPASNTKTEINTKTS QKOTSLKAAPAGSTYNSLPDPDNLAKLAVIITGNAATGNEVDSALALISQDL LSGEINDPDISINISBQYLENLISLSENNISDIAPKDLVNLVLSNRLV NLSGVEDLVNLQELNVSANKALEDISQVSLPVLPKEISAGCNIKTELPNAGAVP ELKETPYLQENDLTNLTLAKLPYKKNLYIKGNLSLSLETNGATKQOLIDANCTDL ETLTGISLSELEMIQLSGCSKLEIKTSLKNPLNVLITADSCAIBDLGTINLPLKO TLVLSNENLTNTAITDLPKLTITLDGGGITSIGTLNLPLEKLDIKENOITIS EITDLPRLSYDVSNNLTITDLPKLTILEMILNVSNRSLSDVSTINPESLYINIS NVIRTVGKTELPSEKERYQNNISDSISMTIDWPKRVKVDASNNLITNIGFDDP KQSLDVHSNRITSTVTHDLPSELETFAQNTLITNIGMDNLPLTYVNLSENRIPS LAETGDLPNLETIVSDNNISYKSLGTMDGVPKRLIDLQNNYLNANTYEGNLSLSD LTNITELNENNYIDDISGLSTSLRLIYLNLSNKLIEDISALSNLTNLQELLENKK IENISALSDLENKLVSKKLIIDISPVANNRGAIVTASNOTYTITPVTYSYSSP TIDNPVTWYDGTLLABSSIGNSGNYDGTITYNMRTSSSTIPNRRKLDGCTFSGT VTDPYKSAKVTADAEQTYTIGDTISBEQPLKDVNAKSSGAAVTSDFATVVDLNTFG

EYVLTITSEKGIQGSCKYIVKVLGAPVISAQDITISYKHATITTEKOFLEDIHAST DLTATITNSTVNNKGGDYVVALNSEMEDGVKATVYVTVNNDPAPIISAKTE IITVDFEKTKEALADIDTADTNDGSVTSNPNFANVLDAGVDTVTLNSINSVGAG TPRAIVYHVEKKEKATISTNTAOQVEYKAINFOPKDVHASINASPPTAVESDPE TVYKLVDPGTYTITATNTNDGVSNAKESVIVRKIPAEITADKEITTEKPREVSE ABRLNDIHATISKKNVAITSNFTDVNLNAGDTVTLNATNEGVKATPEVAVHVO QGEPVITATISTYKPNITAKFLEDIHATNSDQSSVITSNFOTATNFKTAMS YTVTLNANVEDGISABPVAVTIVNKEPALAKADEVSYAKNAVTEPFPKHLE GTAPASAKATSNDSVDRSKTGDTVTITATNDEDAVSTPIEIVHIEBSAPVT ANAEVKNKHEOTDERPFLYDSRKAIDEANVBIKTDEAEKVDINKVGTVYVTLTATNE DGAANPVERSVYSDAAAKNAVKYVDENGSEISAETITGNIDEPESIDAKSIAGY KCAITISGVSTIVEQVVFHAKIKGVVITIKEDTNGKAVDEKQITGEVDGDFEAB AQVDSYSCRAIASGKITEBPQITITFTYSTATPSKSGEITVOYVDSGKGLASKV TGNIDDSYVEAKAIBGSYVGDDSAKGVTEKSGQVTFPKYKNTQVSQDDPKYKGT NOSSSTDTKLVNNNSLIPATGTEMMILAVLIGFNNMLIVASIFLFRPKXTNQ"

6491. 7032  
/gene="lmo0334"  
6491. 6496  
/gene="lmo0334"  
6508. 7032  
/gene="lmo0334"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC98413.1"  
/db\_xref="GI:16409712"  
/db\_xref="UniProt/TrEMBL:O8YA31"  
/translation="MKKTIYSQSGYATNELSOKAIPMYSVESIPESIVAE IYVGSIVMGKVTGDPKSKKYTLSSKLIVISAGVDPKPKENTNDIETLVENLQK TNFLKDIFCKRKLDTPQDLRLKHKLIVNALYSAKAKKESDLNDRDITVRAVEDDS QIDLNKLDGVLDRL"  
join(7235..7240,7248..7532)  
7235..7240  
/gene="lmo0335"  
7248..7532  
/gene="lmo0335"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC98414.1"  
/db\_xref="GI:16409713"  
/db\_xref="InterPro:IPR002453"  
/db\_xref="UniProt/TrEMBL:O8YA30"  
/translation="MREIVDERTFQOHTKLSBSSTGSLPLKNGMAYSKANSIDO LBSALIELVVDVEDFOHTVKDASRLKKGIAVAKODQLMGKXINQLEVR"  
7519..7524  
/gene="lmo0336"  
7519..7524  
/gene="lmo0336"  
7533..7901  
/gene="lmo0336"  
7533..7901  
/gene="lmo0336"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC98415.1"  
/db\_xref="GI:16409714"  
/db\_xref="UniProt/TrEMBL:O8YA29"  
/translation="MDSASLQECQIQIKKAKITAFADQTRKRONBDMLEPDSN QUTQERVILDLPSGMOSEASFSFRITLEBOHESQVWRDLQDKRSDLDTLEQEN KQRLHLEKAKQATLQKEMSQ"  
7885..7890  
/gene="lmo0337"  
7885..7890  
/gene="lmo0337"  
7898..9373  
/gene="lmo0337"  
7898..9373  
/gene="lmo0337"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAC98416.1"  
/db\_xref="GI:16409715"  
/db\_xref="InterPro:IPR010676"

```

/db_xref="UniProt/TREMBL:Q8YA28"
/translation="MSRIDIAELNDPLHGLSSNAEKAMIRIKIEKAMADYADONRLK
GEAVSTKQVSTSTYTSICOSIIEALDESEBIAOYIREFSGOVSSPAAIDIEILQ
EAMAKVGLQKREEDLHROLTAHPMTKMOOVAVKSGSVTOLKIMIEGNITKELT
AFESHGQFSALEBELIATRAVAOELLHHHSFNDKCTYVPSAANSLIMKARD
KARENDKDPKPAFEDYVAIVYVNDQSTVTMUEKQKVENLEODPLEKKG
QELDPLLYTNLSGELERKNDSWKEGINLNGQKVSQSGATLRSAYVASMEDWD
DAGLDMALGLGFGIAARANKPVIKEKLNPAETKQTVNGEKMNNYFKKYGAVNOMK
PSPEDDIIINPERLYGSTKKEVROMLGEKMEETYSAGNGMKFTNEGDVAFYHPGE
GIHKGSYGFSGSDPTKVKIYGVKQYIDFSKDKATIIKRGGE"
gene
9362..9368
/gene="Imo0338"
RBS
9362..9368
/gene="Imo0338"
gene
9377..9757
/gene="Imo0338"
CDS
9377..9757
/gene="Imo0338"
/codon_start=1
/protein_id="CAC98417.1"
/db_xref="GI:16409716"
/translation="WKKLTVGGKFDWYIDLEKSGMFLIEISNEIETPEDPIYGV
TSFSPKNNLSLSEKANEIIDEEMEKNAVILREKVLKIDMTDILNINSVRSSEWLDIFR
RDELKROLKERMSDEIEYLKTL"
gene
10019..10024
/gene="Imo0339"
CDS
10019..10024
/gene="Imo0339"
/translation="weakly similar to inorganic pyrophosphatase"
/codon_start=1
/protein_id="CAC98418.1"
/db_xref="GI:16409717"
/db_xref="GOA:Q8YA26"
/db_xref="InterPro:IPR008162"
/db_xref="InterPro:IPR008163"
/db_xref="UniProt/TREMBL:Q8YA26"

```

Query Match 23.5%; Score 233.6; DB 1; Length 280050;  
 Best Local Similarity 53.1%; Pred. No. 3e-52;  
 Matches 497; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

```

3 TATGTCACAACTCTTACCTTGAAGAACTCTGACCTTAACATGTTATTCAGAAACAT 62
118720 TATGTCACAACTCTGTTTGAAGAACTGTAAGTGAAGAAACATTTATTCAGAAACAT 118779
63 GGATTTGATTTGAGTGGGAGACAGAGGTGATTCCTATCCGCCGCTTACAGCTGGA 122
118780 GGATTTGCTTTATTTTGAAGCGCAACCCACAACTTACCCAGAAATTTATGANTGGA 118839
123 CAGTGAAGTGAAGGAGGCCCATCAGACACAGTACGGCTTATTCGTTATGCGTATGCGAGAA 182
118840 ATCTCTACAGACGGGCAATATATATATGATATGATTTGTTGAGCGGAGAGAGA 118899
183 ACTTGAATATATATTTTGGCCGACGCAATTAAGAACCGTTTATCTTGCGCGCGCT 242
118900 GTTAGATAATATCATTTTGCAGACGGAATGAAGAAAGATTAAGCTGTGACATCTCT 118959
243 TTATTTCCGGCTATGCGAGATACGAAACGATACGGGAAGATACCGTTACATTTGT 302
118960 ATATCTTCAGAGAGAGAGATTTATGCGCCAGACACAGTGAAGAAATCAATTTAGC 119019
303 CCCGATGAGTTTGACATGAGTGTCTGATCTCTCACTCTTGAAGACGCTAAAGAA 362
119020 TCCACAGAGATTTTCTTGTGCTGTAGAACCTTCGCAACGATTAAGATGTGGAAC 119079
363 AAAAGATTCATCTTACGATTTGAGAAATTAATGATCTATGATGATACGTTTACC 422
119080 AAAGTATCGGTATTAACCTGTGATCAACAGTACCGTTTACTTGGAATCACAACACC 119139

```

```

423 GCTTCACTGATATTTGTCAAGCCGAGCGGCCGAACCTGACGATGAAACCAAGACGA 482
119140 ATTAACCTGATTTTCAAGCAAGATGCGCGCTTGATGATGATTTGAACCAAGAAC 119199
483 CCGGCTCAAACTCTATATATCAACCTGCTGTATGACGAGACGCCCGCACTTATATG 542
119200 ATTCCTTCGATTAAGAAACCTTGATGAGATTAAGCAAAATGCGCACGATCGAATG 119259
543 GCATGTAAACCAATCTGACAGCAATATACAGAAATCAGACCGAAGCAATTAAGAGCAAGA 602
119260 GCATTCGAGAAATTTAGCAACTACAGAGTTTACAGCAACACATTAAGCGCGGTAA 119319
603 GATGCGCGGATTAAGCCCTTCTGCTTTGCGCAAGCTTGAAGAACTGTGCTGCCGG 662
119320 GTTCGAGAGATATGCTTAACCATTTTCCAGAGTACAGAAACAGCAAAATTAACCGG 119379
663 GGATTTATACCCGCTTCCGCTTGTGACAGCGTGTATTTGAAGAACATCTAGAGCC 722
119380 TGTATATCTCCGCAAGACGTTTGTCCGCGCATCTTGAAAGAAATATATATAA 119439
723 GCGCGCGATGAACGAAAGGTATACAGCCGCTTTCAATTTTGGCAATATGACGAT 782
119440 AGCCAAACCAAGAAAGAGCTATACAAATGTGTGATGCTTGAACAGTCCGAT 119499
783 ACCAAAGGCGCAGTATACAGAAAGACGAAATTCATTATACGCAATATCTCCGT 842
119500 TCCGAATGAGCGGATTAATTAAGAAAGTGTGACCCGATTTTCAACAAATATGTGGCATC 119559
843 GATGTGCAACGAATCTGAAACTATCTATTTCCACCATATGACATGCGCAATCCAAA 902
119560 TATGTGTCGAAAGCAAAAGCTATCTATTTACCTCATACGAAACCAACCAATCAATAG 119619
903 AGTAAATTTATTTATGAAAGACCTTGACTGTTTGA 938
119620 TGTATCTTGTGCAAGAGATTAAGAAATATGAA 119655

```

RESULT 14  
 AX641665  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Listeria monocytogenes  
 Listeria monocytogenes  
 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 REFERENCES  
 1 Buchrieser, C., Frangeul, L., Couve, E., Ruanlok, C., Fehli, H.,  
 Deloux, P., Dussanget, O., Chetouani, F., Nedjati, H., Glaser, P.,  
 Kunet, F., Cosset, P., Daniels, J., Goebel, W., Kreft, J., Kuhn, M.,  
 Ng, E., Vazquez-Boland, J., Dominguez-Bernal, G., Garrido-Garcia, P.,  
 Trierer-Martinez, A., Amend, A., Chakraborty, T., Domann, E., Hain, T.,  
 Berche, P., Charbit, A., Durant, L., Perez-Diaz, J. C., Baquero, F.,  
 Garcia del Portillo, F., Gomez-Lopez, N., Madueno, E., de Pablos, B.,  
 Wehland, J., Kaerest, U., Entian, K. D., Haut, J., Kose, M., and Voss, H.  
 Listeria monocytogenes genome, plasmids and uses  
 Patent: WO 0101118-A 2655 11-APR-2001;  
 INSTITUT PASTEUR (FR)  
 LOCATION/Qualifiers  
 1..349980  
 /organism="Listeria monocytogenes"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1639"  
 /note="Seq 1 for big: original length: 2944528  
 splitted into 10 <223-seq 0001 : From 0.000.001 to  
 0.349.980 length: 349.980 <223-seq 2855 : From 0.300.001  
 to 0.649.980 length: 349.980 <223-seq 2856 : From  
 0.600.001 to 0.949.980 length: 349.980 <223-seq 2857 :  
 From 0.900.001 to 1.249.980 length: 349.980 <223-seq 2858  
 : From 1.200.001 to 1.549.980 length: 349.980 <223-seq  
 2859 : From 1.500.001 to 1.849.980 length: 349.980

## ORIGIN

<223>seq 2860 : From 1,800,001 to 2,149,980 length: 349,980  
 <223>seq 2861 : From 2,100,001 to 2,449,980 length: 349,980  
 <223>seq 2862 : From 2,400,001 to 2,749,980 length: 349,980  
 <223>seq 2863 : From 2,700,001 to 2,944,528 length: 244,528"

Query Match 23.5%; Score 233.6; DB 6; Length 349980;  
 Best Local Similarity 53.1%; Pred. No. 3e-52;  
 Matches 497; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

```

QY 3 TANGGACAGAGCTTACATTGGAAGAGTGTGACCGCTTAACATCTATTAGCAAGAACAT 62
DB 178720 TATGTGACAGAGCTTGTGATTAAGAAAGTTAGAGAAAACATTTATTATCAAGAACAT 178779
QY 63 GGAATTTGCAATTCAGCTGGGGACAGAGTGAATCTCTATCCGGCCGCTTACAGCTGGA 122
DB 178780 GGAATTTGCTTTATTTTGGAGCGGAACCAACATTAAGCCCAAGAAATTTAGTGGAA 178839
QY 123 CAGTGAAGCTGACGGAAGGGCCCATCAGACAGTACGCGTTATCGGTATGGGGAGAA 182
DB 178840 ATCTTACAGACGGTCCCATATATATTAATAGATATGCAATTTGTTGAGCGGGAGAGA 178899
QY 183 ACTTGAATATATTAATTTCCGACCGCATTAAGAAAGCGTTTATCTTGTGGCGCT 242
DB 178900 GTTAGAATTAATACATTTTCCGACCGATTAAGAAAGATTAAGCTGTGATCTCT 178959
QY 243 TTAATTTCCGGGCTATGCGAGATACGAAACGATACGGGAAGATACGTTACATGT 302
DB 178960 AATCTTCCAGAGAGAGAGATTAATGCGCACACCGTGGAGAAATTAATTAATTTAGC 179019
QY 303 CCGGACGAGTTTGTGACATGAGTGTGCTGCTGCTGCTGCTGCTTGTGGAGACGTAAAGA 362
DB 179020 TCCACAAGATTTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 179079
QY 363 AAAAGATTCATCTTAAACGATTTGAGAGAAAATTAATGATCTATTGGATACAGTTTACC 422
DB 179080 AAAAGTATCGGTATTAACCTTGTGATCAACAGTAAACGCTTATCTGGATCAACACCC 179139
QY 423 GCTTCACTGGAATTTGTCAGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACG 482
DB 179140 ATTACCTGATTTTACAGACAGAGTGGCGCTTGTGATGATTTGACCAACAGAAC 179199
QY 483 CCGGCTCAAAAGTCTATGATATCAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
DB 179200 ATCGCTCGTATTAAGAAAACCTGTAGAGTTATGACAAATACGCAAGCATCGAATG 179259
QY 543 GCATGTAACCAATCTGACAGCAATATACAGAAATCAAGCAATTAAGAGCAAGA 602
DB 179260 GCATGTAACCAATTTACGCAATCTACAGAGTTTACAGCAACAAATTAAGCGCGGTGA 179319
QY 603 GATGGCGGATTAAGCCCTTCTGCTTGTGGCCAAAGGCTTAAGAACTGTGTGCTGCGGG 662
DB 179320 GTTCCGAGATATATGCTTAACCAATTTTCCAAAGGTAACGAAACCAATTAACAGG 179379
QY 663 GGAATTAACACCGCTTCCCGGTTTGTACAGAGCTGTTTATTTGAAAGAACATCTAGAGCC 722
DB 179380 TGGTATATCTCCGCGCAAGAGTTTGTTCGTTGCGGATCTTGAAGAAATTTATTA 179439
QY 723 GGGCGCGGATGAACGAAGAGTGAACAGCGCTTTTCAAAATTTTGGCAATTTAGCGAT 782
DB 179440 ACCGAAACGAAGAAAGCTATCAAAATGTGTGTATGTCTTAAACAGTGTCCGAT 179499
QY 783 ACCAAAGGCGAGTATAGAGAAAGCAAGAAATTCATTAATACCAATTAATCTCCGT 842
DB 179500 TCCGATGAGAGCGGTAATTAAGAAAGTGTGACCTGATTTCAACAAATTTGTGCAATC 179559
QY 843 GATGTGCAAGAACTGAAATCTATTTTCAACCAATGACCAATGACCAATGCGCAATTCAGAA 902
DB 179560 TATGTGTTGCGAAGCAAGCAAGTATTTTCACTATGCAAGAAACCAACCAATTAATG 179619
QY 903 AGTAATTTATTTCAAGAGACCTTGAATCTTTGGA 938
  
```

DB 179620 TGTACTTTGTGCGAAGAACTTATAGAAATATGAA 179655

RESULT 15  
 AX413775  
 LOCUS AX413775 990 bp DNA linear PAT 02-SEP-2002  
 DEFINITION Sequence 766 from Patent WO0228891.  
 ACCESSION AX413775  
 VERSION AX413775.1 GI:21446232  
 KEYWORDS  
 SOURCE Listeria monocytogenes EGD-e  
 ORGANISM Listeria monocytogenes EGD-e.  
 Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 REFERENCE 1  
 Kunst, F. and Glaser, P.  
 Listeria innocua, genome and applications  
 Patent: WO 0228891-A 766 11-APR-2002;  
 INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE  
 SCIENTIFIQUE (CNRS) (FR)  
 Location/Qualifiers  
 1..990  
 /organism="Listeria monocytogenes EGD-e"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:169963"

## ORIGIN

Query Match 23.4%; Score 232.6; DB 6; Length 990;  
 Best Local Similarity 53.0%; Pred. No. 6.1e-52;  
 Matches 496; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

```

QY 4 ATGTGCAAGTCTTACATTGGAAGAGTGTGACCGCTTAACATCTATTAGCAAGAACAT 63
DB 1 ATGTGCAAGCTTGTGATTAAGAAAGTTAGAGAAAACATTTATTATCAAGAACATG 60
QY 64 GATTTGCAATTCAGCTGGGGACAGAGTGAATCTCTATCCGGCCGCTTACAGCTGGA 123
DB 61 GATTTGCTTTATTTTGGAGCGGAACCCAAACATTAAGCCCAAGAAATTAATGATGAA 120
QY 124 AGTGAAGCTGACGGAAGGGCCCATCAGACAGTACGCGTTATCGGTATGGGGAGAA 183
DB 121 TCTTCAAGACGGTCCCAATTTATTAATAGATATGCAATTTGTTGAGCGGGAAGAG 180
QY 184 CTTGAAATATATTAATTTCCGACCGCATTAAGAAAGCGTTATCTTGTGCGGCTT 243
DB 181 TTAGATTAATATCTTTGCAAGCGATTAAGTGAAGAGATTAAGCTGTGATCTTA 240
QY 244 TATTTTCCGGGCTATGCGAGTACGAAACGATACGGGAATACCGTTCAATGTC 303
DB 241 TATCTTCCAGGAAGCAGCTTTATGCGCCAGACCGAGTGAAGAAATTAATTAATGCT 300
QY 304 CCGGATGAGTTTGTGACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
DB 301 CCAAGAAATTTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 364 AAGATTCATCTTAAACATGTAAGAAATTAATGATCTATTGATACAGTTTACCG 423
DB 361 AAGTATCGTATTAACCTTGTGATGATCAACAGTACCGTTACTTGAATCAACACCA 420
QY 424 CTTCACTGATATTTGTCAGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACG 483
DB 421 TTACCTGATTTTTCACAGCAAGAGTGGCGCTTGTGATGATTAAGCAACAGAAACA 480
QY 484 GGCCTCAAAAGTCTATGATATCAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB 481 TGGCTTGTATTAAGAAACCTGTGAGAGTTATGACAAATACGACGATCGAATG 540
QY 544 CATGTAAACATCTGACAGATATACAGGAATCAAGCCAGCAATTAAGAGCAAGAG 603
DB 541 CATATCGAATTTACGCAATCAACAGTTTACAGCAACCAATTAAGCGCGGTGAAG 600
QY 604 ATGGCGGATTAAGCCCTTCTGCTTGTGGCAAGCTTAAGAACTGTGTGCTGCGGG 663
DB 601 TTGGAGAGTATATGCTTAACCAATTTTCCAGAGTACAGAAACAAATTAACAGGT 660
  
```

```

Oy 664 GATTATACACCGCCTTCCCGTTTGTGACAGCTGTTTATTGAAAGAACATCTAGAGCCG 723
    |||||
Db 661 GGTATCTCTCCGCGAGAAAGTTTGTCCGTGCGCATACTTGAAAGAAAATATTAATAA 720
    |||||
Oy 724 GCGGCCGATGAAACGAAAGGTGTAACAGCCGCTTTCAAAATTTGGCAAAATATGACGATA 783
    |||||
Db 721 GCGAAAAACGAAAGAACTATCACAATGTGTGTGTCTTGACAGTGTCCGCAT 780
    |||||
Oy 784 CCAAGGCGCGAGTATTAACGGAAGAACGAAATTCATTATACGCAATATATCTTCCGTG 843
    |||||
Db 781 CCGAATGAGCGGTATATTAAGAAAGTGTGACCTGATTCACACAATATATGTGCACT 840
    |||||
Oy 844 ATGTCCAAAGAACTGGAACCTACTATTTCCACCACTATGACATCGCAATCCAAAA 903
    |||||
Db 841 ATGTGTGGAAGCAAAAGTACTATTTCACTCATACGAAACACCAAAATCAATAGT 900
    |||||
Oy 904 GTAATTTATTTTCATGAAGACCTGACTGTGGA 938
    |||||
Db 901 GTTACTTTGTGAGGAAGTTATAGAAAATATGAA 935
    |||||

```

Search completed: March 19, 2006, 01:01:39  
 Job time : 5418 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:41:32 ; Search time 650 Seconds  
(without alignments)

10212.360 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996  
Sequence: 1 cctaatgtcgaacaaagtcttac.....atgagcttaataagatcc 996

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_21:\*

- 1: geneseq1980s:\*
- 2: geneseq1990s:\*
- 3: geneseq2000s:\*
- 4: geneseq2001as:\*
- 5: geneseq2001bs:\*
- 6: geneseq2002as:\*
- 7: geneseq2002bs:\*
- 8: geneseq2003as:\*
- 9: geneseq2003bs:\*
- 10: geneseq2003cs:\*
- 11: geneseq2003ds:\*
- 12: geneseq2004as:\*
- 13: geneseq2004bs:\*
- 14: geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	996	100.0	996	14 AEA39623	Aea39623 Bacillus
2	439.8	44.2	990	6 ABK74428	Abk74428 Bacillus
3	287.8	28.9	984	8 ACA21250	ACA21250 Prokaryot
4	237	23.8	969	8 ACA21947	ACA21947 Prokaryot
5	233.6	23.5	110000	6 ABA03041_04	Continuation (5 of
6	232.6	23.4	990	6 ABO69869	Abog69869 Listeria
7	232.6	23.4	990	6 ABO67953	Abog67953 Listeria
8	232.6	23.4	990	6 ACA36885	ACA36885 Prokaryot
9	196.4	19.7	1087	13 ADR89833	ADR89833 Clostridi
10	177	17.8	110000	6 ABA90521_18	Continuation (19 o
11	177	17.8	110000	6 ABA90521_19	Continuation (20 o
12	173.8	17.4	1616	12 ACPE57507	ACPE57507 B sphaeri
13	173.8	17.4	1616	12 ADO62865	AD62865 B sphaeri
14	160.2	16.1	978	8 ACA46430	ACA46430 Prokaryot
15	159	16.0	990	4 AAH52510	AAH52510 S. epider
16	159	16.0	996	6 ABN92958	ABN92958 Staphyloc
17	159	16.0	996	13 ADO502160	ADO502160 Staphyloc
18	159	16.0	2975	4 AAH53981	AAH53981 S. epider
19	159	16.0	4114	4 AAH54820	AAH54820 S. epider

20	146	14.7	993	8 ACA20090	ACA20090 Prokaryot
21	146	14.7	999	8 ACF73532	ACF73532 Staphyloc
22	142.8	14.3	1002	4 AAH54746	AAH54746 Staphyloc
23	142.4	14.3	981	4 AAH52083	AAH52083 Staphyloc
24	127.6	12.8	11466	2 AAH74755	AAH74755 Staphyloc
25	117.8	11.8	972	8 ACA47889	ACA47889 Prokaryot
26	106.8	10.7	110000	10 ADF77343_00	ADF77343 Lactic ac
27	104	10.4	4027	13 ADR89831	ADR89831 Lactobaci
28	104	10.4	110000	10 ADF77343_14	Continuation (15 o
29	97.6	9.8	975	13 ADR89827	ADR89827 Lactobaci
30	96	9.6	975	2 AAQ38722	AAQ38722 BSH gene
31	95.4	9.6	972	8 ACA33036	ACA33036 Prokaryot
32	95.4	9.6	987	10 ADH82953	ADH82953 Enterococ
33	82.6	8.3	978	10 ADC90739	ADC90739 E. faeciu
34	80.6	8.1	972	8 ACA33846	ACA33846 Prokaryot
35	80.2	8.1	349980	6 ABO81846	ABO81846 Bifidobac
36	79.4	8.0	110000	10 ADF77343_12	Continuation (13 o
37	79.2	8.0	954	4 AAF89390	AAF89390 Bifidobac
38	72.4	7.3	1050	8 ACA24329	ACA24329 Prokaryot
39	63.6	6.4	978	6 ABO70013	ABO70013 Listeria
40	63.6	6.4	978	6 ABO68098	ABO68098 Listeria
41	63.6	6.4	110000	6 ABA03041_21	Continuation (22 o
42	62.2	6.2	2000	8 ADA71938	ADA71938 Rice gene
43	61.8	6.2	579	8 ACA33669	ACA33669 Prokaryot
44	55.2	5.5	1092	8 ACA52200	ACA52200 Prokaryot
45	54.8	5.5	4118	12 ADO32293	ADO32293 Bacterial

## ALIGNMENTS

RESULT 1	AEA39623	standard; DNA; 996 BP.
ID	AEA39623	
AC	AEA39623;	
XX		
DT	08-SEP-2005 (fixet entry)	
XX		
DE	Bacillus subtilis conjugated bile acid hydrolase gene, SEQ ID NO.1.	
KW	Plasmid, conjugated bile acid hydrolase; penicillin V acylase;	
KW	chologlycine hydrolase; penicillin amldase; gene; ds.	
XX		
OS	Bacillus subtilis; NCIMB 11621.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	4..990
FT		/*tag= a
FT		/product= "Conjugated bile acid hydrolase gene"
FT		/EC number= "3.5.1.24"
FT		/note= "Has penicillin V acylase (EC 3.5.1.11) activity"
XX		
PN	US2005142652-A1.	
XX		
PD	30-JUN-2005.	
XX		
PF	30-MAR-2004; 2004US-00812387.	
XX		
PR	24-DEC-2003; 2003WO-1B006198.	
XX		
PA	(COUL ) COUNCIL SCI & IND RES SOUTH AFRICA.	
XX	(UYUO-) UNIV YORK.	
PI	Sivraman H, Pundle AV, Sureeh CG, Dodson GG, Brannigan JA;	
XX		
DR	WPI; 2005-457790/46.	
XX		
PT	New recombinant plasmid containing a chromosomal DNA fragment from	
PT	Bacillus subtilis encoding conjugated bile acid hydrolase, useful for	
PT	producing large amounts of penicillin V acylase.	
XX		
PS	Claim 1; SEQ ID NO 1; Epp; English.	



CC gene, identifying possible functions of unknown open reading frames and  
 CC monitoring gene copy number variation and stability. Monitoring changes  
 CC in expression of genes may be used to provide a representation of the way  
 CC in which *Bacillus* cells adapt to changes in culture conditions.  
 CC environmental stress or other physiological provocation. Extensive follow  
 CC up characterization is unnecessary, when one spot on an array equals one  
 CC gene or one open reading frame, since sequence information is available.  
 CC This sequence represents a genomic sequence tag (GST) used in the method  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

Sequence 990 BP; 247 A; 214 C; 285 G; 244 T; 0 U; 0 Other;

Query Match 44.2%; Score 439.8; DB 6; Length 990;  
 Best Local Similarity 65.3%; Pred. No. 3.1e-129;  
 Matches 645; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

QY 4 ATGTGCAAGTCTTACATTGGAAGTCTGACCCGTTAACTATTAAGCAAGCAATG 63  
 DB 1 ATGTGTAAGATGATCACTGTGACACTGCGCGGAGCATTTATGGGAGAGAGATG 60  
 QY 64 GATTTGCAATTCAGTGGGAGACAGAGTGAATCTATCCGGCCGTTACACTGGAAC 123  
 DB 61 GATTTGATTTGAGCTTAAACGGGAGGAGTCTTGTGATCCGCTGCTCAAAATGAAA 120  
 QY 124 AGTGAAGCTGACGGAAGGGCCCATCAGACACAGTACGCGTTATCGGTATGGGAGAAA 183  
 DB 121 AGTGAAGCGGAGCGGAGCGAGCATGCCGCGGTATGCGTTATCGGATGGGCAAGAG 180  
 QY 184 CTGGAAATATATATTTTGGCCGACGCGATTAATGAAGCGGTTATCTGTGGCGCTT 243  
 DB 181 CTCAGAAATGCTCTGTTTGGCCGATGATGAATGAAGATTTCTTGGCGGCGCTT 240  
 QY 244 TATTTTCCGGGCTATGCGGAGTACGAAAGAACATACGGAATACCGTTCACTGTTC 303  
 DB 241 TATTTTCCGGGCTATGCTGATGTAAGAAAGAAAGCAAAAGCAAGCCGCAATCTTGGG 300  
 QY 304 CCGCATGATTTGTGACATGCGTGTCTGTCACTGTCTCACTTTGGAGAACGTAAAGAA 363  
 DB 301 CCGCAGCATTCGTGACATGGGTGCTGTGGAATGGCGGATTTGGAGATGTGAAAAA 360  
 QY 364 AAGATTCATCTTTAAGCATTTGAGAAAAAATTAATCATTTGATACAGTTTACCG 423  
 DB 361 GCGGCTGTTCTTTAAACATTTGAGAGAGGTGAGCTTGTGAGTACGCTGAGCGCG 420  
 QY 424 CTTCATGATATTTGTCAACCGGAGCGGCCGAAACCTGACATACCAACAGAGAGC 483  
 DB 421 CTGCATTTGCTATTAACGAGACCGCTCGGGGCAAGCGTGTGTTAAACCGGAGCAGC 480  
 QY 484 GGCCTCAAGTCTATGATATATCAACTGCTGTCATGACAGCAAGCCCGCATTTATATG 543  
 DB 481 GGCATTTCAATTCATGATATATCCGCTCGCGCTGCTGACAGCAAGTCCGATTTCCCTT 540  
 QY 544 CATGTACCAATCTGACAGCATATATACAGAAATCAGCCGAGCAATTAAGAGCAAGAG 603  
 DB 541 CATTTGACCAATTTGGCAACTTTATCGGCTTCAACCGGAGCAGTTTCCCGGAAAAA 600  
 QY 604 ATGGCGGATTAAGCCCTTCTGCTTTTGGCCAAAGGCTTGAAGAACTTGTCTGCGCGGG 663  
 DB 601 ATGGAGGCTGACGCTGCTGCTTTTGTCAAGGTTCCGGCGCTGTCCGATTTGGCAGAG 660  
 QY 664 GATTATACACCGCTTCCCGTTTGTCAAGCTGTTTATTTGAAGAAATCATCAAGCCG 723  
 DB 661 GATTTTACGCGCCCATCCGCTTTGTCAAGGCTGATTTTGAAGAGCAATGAAGCTT 720  
 QY 724 GCGCGCGATGAAGCAAGAGCTGTAAAGCGCTTTCAAAATTTGGCAAAATGAAGATA 783  
 DB 721 GTGTCCGATGAAGCAAGAGCGCTCTCTGCGCGCTTCAATGCTCTCCAAATATGATAT 780  
 QY 784 CCAAGGGGCGAGTATACGGAAGAGCAAAATTCATTAACGAAATATCTTCCG 843  
 DB 781 CTTAAAGGGGTTGTATGACTATCAAGGCGAAAGCAATTAACCCAAATATGCGGGGTA 840

QY 844 ATGTGCAAGAACTGAAACTACTATTTCCACCATATGACATCGCAATCCAAAA 903  
 DB 841 ATGTGAGGAGAGACGCGACCTACTTTTTCACATTATCAAAACAGGAGATTCAAG 900  
 QY 904 GTAATTTATTTTCATGAACCTTGAATGTTTGGAGCCTTAAAGTGTTCGGTAAAGCA 963  
 DB 901 GTTGTCTTATTAACAGAAACCTTGAACGACCAAGATTAAGCGTTTACGCAAGCAG 960  
 QY 964 GAAGAGATTCATGAGCTTAAATTA 990  
 DB 961 GAAGAGCGTTCAATCTCTTCATCA 987

RESULT 3  
 ACA21250  
 ID ACA21250 standard; DNA; 984 BP.

AC ACA21250;  
 DT 19-JUN-2003 (first entry)  
 XX

DE Prokaryotic essential gene #2907.

KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.

OS *Bacillus anthracis*.

PN WO20027183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.

DR P-PSDB; AB017380.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 9120; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 623 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene



CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp://wipo.int/pub/published\_pct\_sequences

XX Sequence 984 BP; 336 A; 165 C; 207 G; 276 T; 0 U; 0 Other;

Query Match 28.9%; Score 287.8; DB 8; Length 984;  
 Best Local Similarity 56.0%; Pred. No. 1.1e-80;  
 Matches 544; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 4 ATGTCACAACTCTTACATTTGAACTGCTGACCTTAAACATGATTTGCAAGCAACATG 63  
 DB 1 ATGTGTACTAGTTTACATTTAGACACAAAACGGTCAGCATCTTTTGCAGAAACGATG 60  
 QY 64 GATTTCGATTTACAGTGGGAGACAGAGTGTCTCTATCCGCGCCGTTACAGCTGAAAC 123  
 DB 61 GACTTCACATTTAGATTAATGATTAAGAGATTAATCAATTCCTCGACATTTACAGTGAAT 120  
 QY 124 AGTGAAGCTGACGGAAGGGCCCATGACACAGTACGCGTTTATGGTATGGGAGAAA 183  
 DB 121 AATATTAACGGGTGAATCATTAATTAAGAAACATGTCAGCGTGGATGGTATTAATCAT 180  
 QY 184 CTGGAATATATTTATTTGCCACGCGCATTAATGAAGGGTTTATCTTGTGCGCGCTT 243  
 DB 181 CAAGGAAGATCATTAATGCGGACGGAATGAAGCAGTATGACATGTCAACACATC 240  
 QY 244 TATTTTCGGGCTATGCGGAGTACGAAAACGATCGGGAAGATTAACCTTACATTTGTC 303  
 DB 241 TATTTTCGAGATTCGCTACTTATGATCAAAAGATTAAGACACAAACGAATTTGGCT 300  
 QY 304 CGCATGATTTGTGACATGGTGTCTGTCAGTCTGTCTGCTTGTGGAAGACGTTAAAGAA 363  
 DB 301 CATTTGATTTTGTAACTTGAAGTGTGACACAAATTAATTTGTCMAAGTTAAAGAAA 360  
 QY 364 AAGATTCGATCTTTAAGCATTTGTAGGAAAAAATTTGATCTTATGATACAGTTTACCG 423  
 DB 361 TCTGTAGATGATTAACCTTTTGTGATTAACCAATTAACGGAATTTAGACTTAAGCCACA 420  
 QY 424 CTTCACGTGATTTGTGACACCGGACCGGACCGCAACCTGACATAGAACCAAGACGAC 483  
 DB 421 CTACATGATTTTATAGCGATTAATGGGAGATTTGATGATGATCGACAAAGTGA 480  
 QY 484 GGCCTCAAGTCTATGATTAATCAACCTGTGTGATGACGAAACGCCCCGACTTATATGG 543  
 DB 481 GGATTTAAATTTGATGATTAACCACTAGAGATGATACCAATGTGTCGGAGTTTAAATGG 540  
 QY 544 CATGTACCAATCTGACGATTAATACAGAAATACGACCAAGCAATTAAGAGCAAGAG 603  
 DB 541 CATTTTCAAAATTTTAAAGACAAATTAATGAGCTTAAATTCGAGCATTCGGCCAAAGAG 600  
 QY 604 ATGGGCGGATTAAGCCCTTTCTGTGCGCAAGGCTTGAAGACCTTGTGTCGCGGAG 663  
 DB 601 TGGAGTAAATTTACATTAATGCTGTTTGGCAAGGCTCGGGGCTCAATGGGACTTCAAGG 660  
 QY 664 GATTATACACCGCTTCCGCTTGTGACAGCTGTTTATTTGAAGAAACATCTAGACCGG 723  
 DB 661 GATTTACACCGCTCATGAGTGTGTCGGGCGACATATGGCAAAACAAATTCAGAGT 720  
 QY 724 GCGGCGGATGAAGCAAGGTGTAACAGCGCTTTCATTTTGGAAATATGACGATA 783  
 DB 721 ATAGATAGCAAAAGGAGATGATGACCCCTTTTCATATCTTATCAATTTGTAGGCTT 780  
 QY 784 CCAAGGGCGCACTGATTAACGGAAGAAACGAATTCATTAACGCAATTAATCTTCGGT 843

DB 781 CCTTAAAGTGAAGTAAATTAACGAAGAGTGCATTAAGTATATCAATATATCAAGCTTA 840  
 QY 844 ATGTCACAACTGGAACCTGAACTACTATTTCACACATATGACATCGCAATTCAAA 903  
 DB 841 ATGTGTAGATTCGGAACATTAATTAATCAATCTTAATGATTTAAGCAATATATGCT 900  
 QY 904 GTAAATTTATTTCAATGAAGACTTGAAGCTTTTGGAGCTTAAAGTGTTCGCTTAAAGCA 963  
 DB 901 GTTATTTATTTATTAATGAAGAAATTTATGATGATGATGATTAAGCTTATTCGCTCAACGG 960  
 QY 964 GAAGAGATAT 974  
 DB 961 AAACAAAAAT 971

## RESULT 4

ACAA21947  
 ID ACA21947 standard; DNA; 969 BP.

ACAA21947;

19-JUN-2003 (first entry)

DE Prokaryotic essential gene #3604.

KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

OS Bacillus anthracis.

PN WO20027183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US0009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR P-PSDB; ABU18077.

PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 9817; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an

organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acid are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 969 BP; 320 A; 175 C; 206 G; 268 T; 0 U; 0 Other;

Query Match 23.8%; Score 237; DB 8; Length 969;

Best Local Similarity 55.5%; Pred. No. 1.8e-64;

Matches 538; Conservative 0; Mismatches 425; Indels 6; Gaps 4;

```

QY 4 ATGTCACAAGTCTTACATTGGAACCTGACCGTAAACATGATATGACAGAACATG 63
DB 1 ATGTGTAAGTGTGATGACATTAAGACAAAAACGTCAGATCTTTGGCAAGACATG 60
QY 64 GATTTTGCAATTCAGCTGGGACAGAGGTATCTTATCCGCGCCGCTTACAGCTGAC 123
DB 61 GACTTCACATTAGATGATGATCAAGAGTAAATATCATCTCTGACATTCAGTGAAT 120
QY 124 AGTGAAGCTGACGGAAGGCGCCATCAGACACAGTACCGCTTTATCGGTATGGGAGAAA 183
DB 121 AATATACCGGTAAATCATTAATACGAATGCTACGCTGCGAATGGGTATTAATCAT 180
QY 184 CTGGAATATATATATTTATTTGCGACGCGATTAATGAAGGCTTATTTGGGGGCTT 243
DB 181 CAAGGAAGATCATTTATGCGGACGAGTAAATGACAGATATGACATGACACATC 240
QY 244 TATTTTCCGGGCTATGCGGAGTACGAAAAACATACGGAAGATCCGTTCACTTGTG 303
DB 241 TATTTTCAAGATTCGCTACTTATATGCAAGATGATGACAAACAGAAATTTGGCT 300
QY 304 CCGCATGAGTTTGTGACATGGGTGCTCTGCTGCACTTTT--GGAAGAGTAAAG 361
DB 301 CCAATTGATTTTGTATCTTGAAGTGTGACCAATTCATTTCTGCAAGCCAGTGAAG 360
QY 362 AAAAAGTTCGATCTTTAATGATTTGAGAAAAAATGATCTATTGATAC--AGTTTAA 420
DB 361 AATCTGTAGATGACATTTACTTTTGGATATACCAATTCACCGCATTTAGACTTACGCCA 420
QY 421 CCGCTTCACTGATATTTGTCAGACCGGACGCGCCGAAACCTGACATAGAACCAAGACA 480
DB 421 CCACTACATTTGATTTTAAAGGATTAATGGGAGATTTGATTTGATTCGATCCGACAGT 480
QY 481 GACGCGCTCAAGTCTATGATTAATCAACTGTGTGTCATGACGAACGCCGACTTATA 540
DB 481 GAAGGATTAATTAATGATGATTAACCACTGAGTGTGACGAATGTGCGGAGTTTAA 540
QY 541 TGGCATGTAACCAATCTGACGCAATTAACGAAATGACACCGAAGCAATTAAGAGCAAA 600
DB 541 TGGCATTTACAAATTTAAGACATATATAGGCTTAAATGAGCATTGCGGCCAACA 600
QY 601 GAGATGGGCGATTAAGCCCTTTCTGCTTTTGGCCAAAGGCTTAGAAGCTGTGTGCGG 660
DB 601 GAGTGAAGTAATTAACATTAAGTGTGCTTTGGCCAAAGGCTCGGCGCTCAATGGACATTTCCA 660
QY 661 GGGGATTAATACCCGCTTCCGCTTTGTCAGAGCTGTTTATTTGAAAGCAATCTAAG 720
DB 661 GGGGATTTTACCCCGCATGAGGTTTGGCGGCGCAATATGGCAAAACAAACATTTCA 720
QY 721 CCGGCGCGCGATGAAGCAAGGTATACAGCGGCTTTCAATTTTGGCAATATAGAG 780
DB 721 GGTATATGATAGCAAGAAAGGAGTATACGCCCTTTTCAATTTTCAATTAATGTGAG 780

```

```

QY 781 ATACCAAGGCGCAGTATGATGAGAAAGACGAATTCATTTATGACATATA-CTTC 839
DB 781 GTTCTTAAGTGAAGTAATTAAGCAAGACAGATGATATATACATATATACCAAG 840
QY 840 CGTATGTGACAGAACTGGAACCTATCTATTTTCCACCACTATGACAAATCGG--CAATC 897
DB 841 CGTAAATGTGATGAAATCCCGAACATATTTATCTTACTTCAATAGTATGACCAATTA 900
QY 898 CAAAAGTAATTTATTTATGTAAGACCTTGAAGCTGTTGAGGCTTAAGTGTTCGCT 957
DB 901 TAGCTGTATTTATTTATTTCTGAAAATTTATGATACGATGAAATTAAGCTTATCCTTC 960
QY 958 AAAGCAGAA 966
DB 961 CACCCGGA 969

```

# RESULT 5

ABA03041\_04  
Continuation (5 of 30) of ABA03041 from base 400001 (Listeria monocytogenes BGD-e gen  
WP Sequence split into 30 fragments LOCUS ABA03041 accession ABA03041

Fragment Name	Begin	End
ABA03041_00	1	110000
ABA03041_01	100001	210000
ABA03041_02	200001	310000
ABA03041_03	300001	410000
ABA03041_04	400001	510000
ABA03041_05	500001	610000
ABA03041_06	600001	710000
ABA03041_07	700001	810000
ABA03041_08	800001	910000
ABA03041_09	900001	1010000
ABA03041_10	1000001	1110000
ABA03041_11	1100001	1210000
ABA03041_12	1200001	1310000
ABA03041_13	1300001	1410000
ABA03041_14	1400001	1510000
ABA03041_15	1500001	1610000
ABA03041_16	1600001	1710000
ABA03041_17	1700001	1810000
ABA03041_18	1800001	1910000
ABA03041_19	1900001	2010000
ABA03041_20	2000001	2110000
ABA03041_21	2100001	2210000
ABA03041_22	2200001	2310000
ABA03041_23	2300001	2410000
ABA03041_24	2400001	2510000
ABA03041_25	2500001	2610000
ABA03041_26	2600001	2710000
ABA03041_27	2700001	2810000
ABA03041_28	2800001	2910000
ABA03041_29	2900001	2944528

Query Match 23.5%; Score 233.6; DB 6; Length 110000;

Best Local Similarity 53.1%; Pred. No. 2.7e-62;

Matches 497; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

```

QY 3 TATGTCACAAATCTTACATTGGAACCTGCTGACCGTAAACATGATTTGCAAGAACAT 62
DB 78720 TATGTCACAAACCTTCTGATTTGAAAACGTTAGATGAAAACATTTATTAAGAACAT 78779
QY 63 GGAATTTGCAATTCAGCTGGGACAGAGGTATCTTATCCGCGCGCTTACAGCTGGA 122
DB 78780 GGAATTTGCTTTATTTTGAAGGCCAACCAACATTAAGCCCAAGAAATTAAGATGAA 78839
QY 123 CAGTAAGCTGACGGAAGGCGCCATCAGACAGATACGCGTTATCGTATGGGAGAA 182
DB 78840 ATCTTCTACAGCGGTGCATATTAATTAATGATATGCAATTTGTTGAGCGGAGAGAA 78899
QY 183 ACTTGAATATATATTTATTTGCGAGCGCATTAATGAAGCGTTATCTTTGCGCGCT 242
DB 78900 GTTAGATTAATACATTTTTCGACAGCGATGATGATGAAGAGATTAAGCTGTGATCTT 78959

```

QY	243	TTATTTTCCGGGCTATGCGGAGATACGAAAAACGATACGGGAAAGATACGGTTCACATGTG	302
Db	78960	ATATCTTCCAGGAGAAACGATTATATGCGCCAGACCAATGGGAAAGGAAAAATCAATTTAGC	79019
QY	303	CCCGCATGAGTTGTGTCATGAGGAGCTGTCACTGTCTGAGTCTTTTGGAGACGATAAAGA	362
Db	79020	TCCACAAAGAAATTTTGTCTTTGGCTGTTAGGAACTTGGCGAACGATTAAAGATGTGGAGC	79079
QY	363	AAAGATTCGATCTTTAACGATTGTAGAGAAAAAATTAGATCTATTGGATACAGTTTACC	422
Db	79080	AAAGTTATCGGTATTATAACTTGGGATCAACACAGTACCGTTTACTTGGATACACAACCC	79139
QY	423	GCTTCACTGATATTTGTCAAGCCGGACGGCCGAAACCTGACATATGAACCAAGACGAC	482
Db	79140	ATTACACTGATTTTTCACAGACAAAGATGGGCGCTGTGTAGTATTTGAACCAACAGAAAC	79199
QY	483	CGGCTCAAAAGTCTATGATATATCAACTGTGTCTATGACGAACGCCCGGACTTATATG	542
Db	79200	ATCGCTTCGTATTAAGAAAAACCTGTGAGAGTTATGACAAATACGCACAGCATCGAATG	79259
QY	543	GCAATGTAACCAATCTGCAGCAATATACAGGAATCAGACCAGAACATTTAGAGACAAAGA	602
Db	79260	GCAATATGAGAAATTTTACGAACTACACAGTTTATCAAGCAACCAATTTAGCCGCGTGA	79319
QY	603	GATGGCGGATTAAGCCCTTCTGCTTTTGGCCAAAGCTTAGAAGCTGTGTGCTGCCGG	662
Db	79320	GTTTCGAGAGATATATGCTTAAACCATTTTGCAGAGTACAGGAAACCAAGCAATTTACAG	79379
QY	663	GGATTAATACACCGGCTTCCGGTTTGTTCAGAGCTGTTATTTGAAAGAACATTTACAGCC	722
Db	79380	TGTTTATATCTCCGCGAGAACGTTTGTCTGCGGCACTACTTAAAGAAAAATTTATPAA	79439
QY	723	GGCGGCGATGAAGAACGAAGGTGTATCAGCCGCTTTCAATTTTGGCAAAATATGACAT	782
Db	79440	AGCCAAAAACGAAGAAAGAGTATCACAAAATGTGTGTATGTCTTGAAACAGTCCGCAAT	79499
QY	783	ACCAAGGCGCGACATGATTAACGAAGAAAGACGAATTCATTATACGATATATCTTCGAT	842
Db	79500	TCCGATATGAGACGGTATTTAAAGAAAGGTGTGACCTTGATTTACACAAATATGTGGATTC	79559
QY	843	GATGTGCAAGAACTGGAAACTGTCTTTTTCACACATATATGACATGTGGAAATCCAAA	902
Db	79560	TATGTGTTCCGAAAGCAAAAGTACTCTTTTTCACCTCTATACGAAAAACAACCAATCATATG	79619
QY	903	AGTAATTTATTTCAATGAAGACCTGACTGTTTGA	938
Db	79620	TGTTACTTGTTCGAAGAGATTTATGAAATATATATA	79655

RESULT 6	
ABQ69869	
ID	ABQ69869 standard; DNA; 990 BP.
XX	
AC	ABQ69869;
XX	
DT	29-AUG-2003 (revised)
DT	29-AUG-2002 (first entry)
XX	
DE	Listeria monocytogenes EGDe DNA sequence #81.
XX	
KM	Antibacterial; Listeria; food contamination; mutational analysis.
KM	Infection; ds.
XX	
OS	Listeria monocytogenes; EGDe.
XX	
PN	WO200228891-A2.
PD	
XX	11-APR-2002.
XX	
PF	04-OCT-2001; 2001MO-FR003061.
XX	
PR	04-OCT-2000; 2000FR-00012697.

XX (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX Kunst F, Glaser P;  
PI  
DR WPI; 2002-332479/37.  
XX  
XX New genomic sequences from *Listeria* species, useful for detection,  
PT treatment and prevention of infection, also related polypeptides,  
PT antibodies and modulators.  
XX  
XX Claim 16; SEQ ID NO 2682; 180pp; French.  
XX  
XX The present invention relates to nucleic acid sequences (AB067188-  
CC AB071212) from *Listeria* sp. The sequences are useful as probes and  
CC primers for identification and/or detection of *Listeria* (e.g. as  
CC contaminants in foods, or mutational analysis) and for analysis of gene  
CC expression. Proteins encoded by the nucleic acid sequences can be used to  
CC screen for compounds that modulate gene expression, replication and  
CC pathogenicity of *Listeria* (potential therapeutic agents), also for  
CC treating infections by *Listeria*, and are useful as immunogens in anti-  
CC *Listeria* vaccines. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences). (updated  
CC on 29-AUG-2003 to standardise OS field)  
XX  
XX Sequence 990 BP; 347 A; 171 C; 215 G; 257 T; 0 U; 0 Other;

Query Match	Similarity	23.4%	Score 232.6	DB: 63	Length 990
Beet Local	Similarity	53.0%	Pred. No. 4.7e-63		
Match	Conservative	0	Mismatches 439	Indels	Gaps 0
4	ATGTGCACAACTCTTACATTGGAAA	CTGCTGACCGGTAAACATGTATTAGCAAGAACATG	63		
1	ATGTGCACAAAGCTTGCTGTATTGAAAC	CGTTATATGGAAGAAACATTTATTATTCAGAAACGATG	60		
64	GATTTTGCATTTCACTGGGGGACAGAGGT	ATCTCTATCCGCGCGCTTACAGCTGGAAAC	123		
61	GATTTTGTCTTTATTTTGGAGGGCGAAC	CCCAACATTTAGCCCAAGAAATTTATGATGGAAA	120		
124	AGTGAAGCTGACGGAAAGGGCCATCAG	ACACAGTACGGCTTTATCGTATATGGGAGAAA	183		
121	TCTTTTACAGACGGGCGCCATTTATAT	TATATGATATGCAATTTGTGTGGAGCGGAGAGAG	180		
184	CTTGGAAATATATTATTTTGGCCGAC	GGCATTTATATGAAGCGGTTTATCTTTGGCGCGCTT	243		
181	TTAGTAATATACATTTTGGCAGAC	CGGAATGTGATGAAGAAAGGATTTAAGCTGTGCATCTCTA	240		
244	TATTTTCCGGCGTATGCGGAGTACGAAA	AAAAACGATACGGGAAAGATACCGTTACATTTGTC	303		
241	TATCTTCCAGAGAAAGCAAGTTTATCG	CCACGACACCATGTGAAGGAAAAATCAATTTAGCT	300		
304	CCGCACTGAGTTTGTGACATGGGTGCT	GTCACTGTCACTGTCACTCTTTGGAAAGCGTAAAGAA	363		
301	CCACAGAAATTTTCTCTTTGGCTGT	TTAGGAACCTTCGCAACGATTTAAAGATGTGGAAGCA	360		
364	AAGATTCGATCTTTAACGATTTGTAG	AGAAAAATTTAGATCTTATTTGGATACATTTTACGG	423		
361	AAGTTATCGGTATTAACCTTGTGTGAT	CAACGATACCGTTATCTTGGAAATCACACACCA	420		
424	CTTCACTGGATTTATGTCCAGACCGGA	CGGGCCGAAACCTGACGATAGAACCAAGAGCAGAC	483		
421	TTACACTGGAATTTTACAGACCAAGAG	TGGGCGCTTGTATGTGATTTGAACCAACAGAAACA	480		
484	GGCCTCAAAAGTCTATGATATATCAAC	CTGTGGTCTATGACGAAACAGCCCGCACTTTATATGG	543		
481	TGCGTTTCGTATTAAGAAAGAAAC	CCCTGTAGGAGTTATGACAAATACGCAACGATCGAATGG	540		
544	CATGTAAACCAATCTGCAGACATATAC	AGGAATTCAGCCGAGCAATTTAGAGCAAGAG	603		
541	CATATCGAANAATTTACGCAATATCAC	ATGAGTTTACAGACCAACATTTAGCCCGCGGTGAAG	600		

QY 604 ATGGCGGATTAAGCCCTTTCCTTTGGCCAAAGCTTAGAAGTGTGGTCCCGGG 663  
 DB 601 TTGGAGAGTATATGCTTAAACATTTTCGAAAGTACAGGAACACCAATTAACAGGT 660  
 QY 664 GATTATACACCGCTTCCGGTGTTCAGAGCTTTATTTGAAAGACATCTAGACCG 723  
 DB 661 GGTATATCTCGCGCAGAACGTTTTGTCTCGTCCGCAATCTTGAAAGAAATATTTATATA 720  
 QY 724 GCGGCGCATGAAACGAAAGGTGTACAGCCGCTTTCAAAATTTTGCAAAATATAGAGATA 783  
 DB 721 GCGAAACAAAGAGAGCTATCAAAATGTGTATGTCTTGAACAGTGTCCGATT 780  
 QY 784 CCAAGGGCGCAGATGAAGAGAGAGAGAAATTCATTAACGAAATATCTTCCGTG 843  
 DB 781 CCGAATGAGCGGATTAATTAAGAAAGTGTGACCTTATTTCAACAAATATGTGGCATCT 840  
 QY 844 ATGTGCAACGAAACTGGAACACTATTTCCACACATATGACAAATCGCAAAATCCAAA 903  
 DB 841 ATGTGTTCCGAAAGCAAAAGCTATCTTTCACTCATACGAAACAAATCAATATGT 900  
 QY 904 GTAAATTTATTTCAATGAAAGCTTGAAGCTTTGGA 938  
 DB 901 GTTACTTTGTGCAAGAGATTAAGAAATATATGAA 935  
 RESULT 7  
 ABQ67953  
 ID ABQ67953 standard; DNA; 990 BP.  
 XX ABQ67953;  
 AC 29-AUG-2003 (revised)  
 DT 29-AUG-2002 (first entry)  
 DE Listeria monocytogenes EGD DNA sequence #77.  
 XX Listeria monocytogenes EGD DNA sequence #77.  
 KM Antibacterial; Listeria; food contamination; mutational analysis;  
 KM infection; ds.  
 OS Listeria monocytogenes; EGD.  
 XX MO200228891-A2.  
 PN 11-APR-2002.  
 PD 04-OCT-2001; 2001WO-FR003061.  
 XX 04-OCT-2001; 2001WO-FR003061.  
 PF 04-OCT-2000; 2000FR-00012697.  
 PR (INSP) INST PASTEUR.  
 PA (CNRS) CNRS CENT NAT RECH SCI.  
 XX Kunat F, Glaeser P;  
 PI WPI, 2002-332479/37.  
 DR New genomic sequences from Listeria species, useful for detection,  
 PT treatment and prevention of infection, also related polypeptides,  
 PT antibodies and modulators.  
 XX Claim 16; SEQ ID NO 766; 1800p; French.  
 PS The present invention relates to nucleic acid sequences (ABQ67188-  
 CC ABQ71212) from Listeria sp. The sequences are useful as probes and  
 CC primers for identification and/or detection of Listeria (e.g. as  
 CC contaminants in foods, or mutational analysis) and for analysis of gene  
 CC expression. Proteins encoded by the nucleic acid sequences can be used to  
 CC screen for compounds that modulate gene expression, replication and  
 CC pathogenicity of Listeria (potential therapeutic agents), also for  
 CC treating infections by Listeria, and are useful as immunogens in anti-  
 CC Listeria vaccines. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. (Updated

CC on 29-AUG-2003 to standardise OS field)  
 XX SQ Sequence 990 BP; 347 A; 171 C; 215 G; 257 T; 0 U; 0 Other;  
 Query Match 23.4%; Score 232.6; DB 6; Length 990;  
 Best Local Similarity 53.0%; Pred. No. 4,7e-63;  
 Matches 496; Conservative 0; Mismatches 439; Indels 0; Gaps 0;  
 QY 4 ATGTGCAACAGTCTTCAATTTGAAAGTGTGACCTGTAACATGATATAGCAAGAACATG 63  
 DB 1 ATGTGCAACAGTCTTCAATTTGAAAGTGTGACCTGTAACATGATATAGCAAGAACATG 60  
 QY 64 GATTTGCATTTCAGCTGGAGACAGAGTGATTTCTATCCGCGCGTTACAGCTGAC 123  
 DB 61 GATTTGCATTTCATTTTGAGGGGCAACCCACAAATTAAGCCCAAGAAATATATGATGAAA 120  
 QY 124 AGTGAAGCTGAGGAGAGGGGCCATGACAGACAGTACGCTTTATGCGTATGGGAGAAA 183  
 DB 121 TCTTCTACAGAGCGGTCCAAATTAATTAATGATATGCAATTTGTTGAGCGGGAAGAG 180  
 QY 184 CTGGAATAATATATATTTTCCGACGCGCATTAATGAAGCGTTTATCTTGTGCGGCGCT 243  
 DB 181 TTAGATTAATATCATTTTCCAGACGCGATTAATGAAGAAATTAAGCTGTGATCTCTA 240  
 QY 244 TATTTTCCGGGCTATGCGGAGTACGAAAAACGATACGGAAAGATACGTTCAATTGTC 303  
 DB 241 TATCTTCCAGAGAAAGCAGTTTATGCGCCAGACCCAGTGAAGAAATCAATTTAGCT 300  
 QY 304 CCGCATGATTTGTGACATGAGGTGCTGTGATCTGTCACTCTTTGGAAGACGTAAGAA 363  
 DB 301 CCACAGAAATTTTGTGTTGGCTGTGAGAACCTTGGCAACATTAAGATGTGGAAGCA 360  
 QY 364 AAGATTCAGTCTTTAAAGATGTGAGAAATAATAGATCTATGATACAGTTTACCG 423  
 DB 361 AAGTATCGGTTATTAACCTTGTGATCAACAGTACCGTTACTTGAATCACAACCA 420  
 QY 424 CTTCACGTGATATGTGACAGCGGACGCGGCAAACTTGACGATGAAACCAAGACAGAC 483  
 DB 421 TTACACTGATTTTCAACAGACAGAGTGGCGCTGTGTGATGATGAACCAACAGAAACA 480  
 QY 484 GGCCTCAAGTCTATGATTAATCACTGCTGTCAATGACAGACCCCGCATTTATATG 543  
 DB 481 TCGCTTCGTATTAAGAAACCTTGTGAGATTAAGCAAAATGCGCACGATCGATATG 540  
 QY 544 CATGTAAACCATCTCGACGATATATACAGGAATCAAGCCAGCAATTAAGAGCAAGAG 603  
 DB 541 CATATCGAAATTTACGCACTACACAGGTTTCAAGCAACAAATTAAGCGCGCGTGAAG 600  
 QY 604 ATGGCGGATTAAGCCCTTTCCTTTGGCCAAAGCTTAGAAGTGTGGTCCCGGG 663  
 DB 601 TTGGAGAGTATATGCTTAAACATTTTCGAAAGTACAGGAACACCAATTAACAGGT 660  
 QY 664 GATTATACACCGCTTCCGGTGTTCAGAGCTTTATTTGAAAGACATCTAGACCG 723  
 DB 661 GGTATATCTCGCGCAGAACGTTTTGTCTCGTCCGCAATCTTGAAAGAAATATTTATATA 720  
 QY 724 GCGGCGCATGAAACGAAAGGTGTACAGCCGCTTTCAAAATTTTGCAAAATATAGAGATA 783  
 DB 721 GCGAAACAAAGAGAGCTATCAAAATGTGTATGTCTTGAACAGTGTCCGATT 780  
 QY 784 CCAAGGGCGCAGATGAAGAGAGAGAGAAATTCATTAACGAAATATCTTCCGTG 843  
 DB 781 CCGAATGAGCGGATTAATTAAGAAAGTGTGACCTTATTTCAACAAATATGTGGCATCT 840  
 QY 844 ATGTGCAACGAAACTGGAACACTATTTCCACACATATGACAAATCGCAAAATCCAAA 903  
 DB 841 ATGTGTTCCGAAAGCAAAAGCTATCTTTCACTCATACGAAACAAATCAATATGT 900  
 QY 904 GTAAATTTATTTCAATGAAAGCTTGAAGCTTTGGA 938  
 DB 901 GTTACTTTGTGCAAGAGATTAAGAAATATATGAA 935

RESULT 8  
 ID ACA36885  
 ACACA36885 standard; DNA; 990 BP.  
 ACACA36885;  
 19-JUN-2003 (first entry)  
 Prokaryotic essential gene #18542.  
 Antisense, ds, prokaryotic essential gene; cell proliferation;  
 drug design; gene.  
 Listeria monocytogenes.  
 WO200277183-A2.  
 03-OCT-2002.  
 21-MAR-2002; 2002WO-US009107.  
 21-MAR-2001; 2001US-00815242.  
 06-SEP-2001; 2001US-00948993.  
 25-OCT-2001; 2001US-0342923P.  
 08-FEB-2002; 2002US-00072851.  
 06-MAR-2002; 2002US-0362699P.  
 (ELIT-) ELITRA PHARM INC.  
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,  
 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 WPI: 2003-029926/02.  
 P-PsDB: ABU33015.  
 New antisense nucleic acids, useful for identifying proteins or screening  
 for homologous nucleic acids required for cellular proliferation to  
 isolate candidate molecules for rational drug discovery programs.  
 Claim 14; SEQ ID NO 24755; 1766bp; English.  
 The invention relates to an isolated nucleic acid comprising any one of  
 the 623 antisense sequences given in the specification where expression  
 of the nucleic acid inhibits proliferation of a cell. Also included are:  
 (1) a vector comprising a promoter operably linked to the nucleic acid  
 encoding a polypeptide whose expression is inhibited by the antisense  
 nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 polypeptide or its fragment whose expression is inhibited by the  
 antisense nucleic acid; (4) an antibody capable of specifically binding  
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 proliferation or the activity of a gene in an operon required for  
 proliferation; (7) identifying a compound that influences the activity of  
 the gene product or that has an activity against a biological pathway  
 required for proliferation, or that inhibits cellular proliferation; (8)  
 identifying a gene required for cellular proliferation or the biological  
 pathway in which a proliferation-regulated gene or its gene product lies  
 or a gene on which the rest compound that inhibits proliferation of an  
 organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 compound's activity; (11) a culture comprising strains in which the gene  
 product is overexpressed or underexpressed; (12) determining the extent  
 to which each of the strains is present in a culture or collection of  
 strains; or (13) identifying the target of a compound that inhibits the  
 proliferation of an organism. The antisense nucleic acids are useful for  
 identifying proteins or screening for homologous nucleic acids required  
 for cellular proliferation to isolate candidate molecules for rational  
 drug discovery programs, or for screening homologous nucleic acids  
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
*K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from Wipro at  
 CC ftp.wipro.int/pub/published\_pct\_sequences  
 XX

Sequence 990 BP; 347 A; 171 C; 215 G; 257 T; 0 U; 0 Other;  
 Query Match 23.4%; Score 232.6; DB 8; Length 990;  
 Best Local Similarity 53.0%; Pred. No. 4.7e-63;  
 Matches 496; Conservative 0; Mismatches 439; Indels 0; Gaps 0;  
 4 ATGTGACAGAGCTTTACATTTGGAAGTGTGACCGCTTAACATGATTTAGCAAGAACATG 63  
 1 ATGTGACAGAGCTTTACATTTGGAAGTGTGACCGCTTAACATGATTTAGCAAGAACATG 60  
 64 GATTTGATTTGACGCTGGGGGACAGAGGTATTCCTATTCGCGCCGCTTACAGCTGGAAC 123  
 61 GATTTGATTTGATTTGGAAGGCAACCCAACTTATGCAATTTAGATGGA 120  
 124 AGTGAAGCTGACGGAAGGCGCCATCAACACAGTACGCTTTATCGTATGGAAGAAA 183  
 121 TCTTCTACAGCGGTGCCAATTAATTAATGATATGATTTGTTGAGCGGAAGAGAG 180  
 184 CTGGAATATATATTTATTTGCGAGCGCATTTATGAAAGCGTTATCTTGTGCGCGCTT 243  
 181 TTAGTAAATATATTTGCGAGCGCATTTATGAAAGCGTTATCTTGTGCGCGCTT 240  
 244 TATTTTCGGGCTATGCGAGATGAAAGAAACGATTAAGGGAAGTACCGTTCACTGTC 303  
 241 TATCTTCCAGAGAGAGCTTTATGCGCAGACACAGTGAAGAAATCAATTTAGCT 300  
 304 CCGCATGATTTGACATGCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
 301 CCAGAGAAATTTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 364 AAGATTCATCTTTACAGATTTGAGAGAAATTAATGATCTATTTGATGATGATTTACG 423  
 361 AAGTATGCTTATTAATCTTGTGATCAACAGTACCGTATCTGATGATCAACACCA 420  
 424 CTTCATGATATTTGACAGCGAGCGGCGGAAACCTGACATGAAACCAAGACAGAC 483  
 421 TTACCTGATTTTTCACAGACAGAGTGGGCTTGTGATGATTTGATGATGATGATGATGAT 480  
 484 GGCCTCAAGTCTATGATATCAACCTGTGTGATGAGCAAGCCGCACTTTATATG 543  
 481 TCGCTGCTATTAAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGATGATGAT 540  
 544 CATGTAACCAATTTGACAGATTTACAGAAATCAAGCCGATTTGATGATGATGATGATGAT 603  
 541 CATATCGAAGATTTGACAGATTTACAGAAATCAAGCCGATTTGATGATGATGATGATGAT 600  
 604 ATGGCGGATTAAGCCCTTCTGCTTTGGCCAGGCTTGAAGATGTTGTGCTGCGCGGG 663  
 601 TTGGAGATATATGCTTAACCATTTTCCAGATGATGATGATGATGATGATGATGATGATGAT 660  
 664 GATTAATACCGGCTTCCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 723  
 661 GGTATATCTCCGCGCAAGAGTTTGTGCTGCGGATCTGGAAGAAATTTATTA 720  
 724 GCGCCGATGAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783  
 721 CGGAAGAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 784 CCAAGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843  
 781 CCGATGAGCGGATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 844 ATGTGCAAGAAATGGAATCTATTTTCCACATGATGATGATGATGATGATGATGATGATGATGAT 903  
 841 ATGTGCAAGAAATGGAATCTATTTTCCACATGATGATGATGATGATGATGATGATGATGATGAT 900  
 904 GTAATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938  
 901 GTTACTTTGCAAGAGATTAATGAATAATGAA 935  
 RESULT 9  
 ADR89833



WP	ABA90521_01	100001	210000
WP	ABA90521_02	200001	310000
WP	ABA90521_03	300001	410000
WP	ABA90521_04	400001	510000
WP	ABA90521_05	500001	610000
WP	ABA90521_06	600001	710000
WP	ABA90521_07	700001	810000
WP	ABA90521_08	800001	910000
WP	ABA90521_09	900001	1010000
WP	ABA90521_10	1000001	1110000
WP	ABA90521_11	1100001	1210000
WP	ABA90521_12	1200001	1310000
WP	ABA90521_13	1300001	1410000
WP	ABA90521_14	1400001	1510000
WP	ABA90521_15	1500001	1610000
WP	ABA90521_16	1600001	1710000
WP	ABA90521_17	1700001	1810000
WP	ABA90521_18	1800001	1910000
WP	ABA90521_19	1900001	2010000
WP	ABA90521_20	2000001	2110000
WP	ABA90521_21	2100001	2210000
WP	ABA90521_22	2200001	2310000
WP	ABA90521_23	2300001	2365589

Query Match 17.8% Score 177; DB 6; Length 110000;  
Best Local Similarity 50.8%; Pred. No. 3.2e-44;  
Matches 504; Conservative 0; Mismatches 480; Indels 9; Gaps 3;

QY	7	TCACAGCTTTACATTGGAACCTGCTACCGTAAACATGATTAGCAAGAACATGAT	66
DB	104062	TGTAGTATTGTTTACACTAGAGATCAAGATTAATTAACATTTCTTCTGTAACATGAT	104003
QY	67	TTTGATTTTCACTGGGACAGAGTGATCTCTATCCGCG---CCGTACAGCTGAAC	123
DB	104002	TTTCAATTAAGAAATGCGGACAAATTTCTTTATTCAGAAACAGAGATAGGATTT	103943
QY	124	AGTGAAGCTGACGGAAGGCGCCATCAGACAGTACGCGTTATCGATGCG---GAGA	180
DB	103942	GCCCACTTAATGAAGAGACATTTGAAACATCTTAATGCTTTGTTAGGATGGGTGCGATG	103883
QY	181	AACTTGAATATATTTATTTGCCAGCGCATTAATGAAGCGGTTATCTTGCGGCG	240
DB	103882	GAAAGGCGGCATCCAGTACTCTTGTGATGATCAAGAAAGGGTTAATGGGTGCGACT	103823
QY	241	CTTTATTTTCCGGGCTATGCGAGATCGAAAAAGATACCGGTACATT	300
DB	103822	TGTATTTTCCAGGCTATGCTGATGATGATGATCAAAAAATCAAAAGGAAAT	103763
QY	301	GTCGCGCATGATTTGTGACATGGGTGCTGCTGATCTGCTTGGAGAGCTTAA	360
DB	103762	TCACTGATATGTTATTTCCAGCGTCTCTCACTCAAGCGATTTAGAGAAATCAT	103703
QY	361	GAAAGATTCGATCTTTAAGATTTGAGAAAAATTAAGATTA---TTGATACATT	417
DB	103702	GATTTATTTGATTAATAATTTGTTATTAATTAATGATTAATCAACCTTAGGCTTACG	103643
QY	418	TTACCGCTTCACTGATTTGTCAGACCGGACCGGCGAAACCTGCGATGAGAACCA	477
DB	103642	CTCTCTTATCAATTTATTTTCTGATCTGATGTCGCAAGTTGATTTATGACCAAG	103583
QY	478	GCAGACGCGCTCAAGCTATGATTAATCAACTGTGTCTGTCAGACGACCGCCGAT	537
DB	103582	CAAGTGAGCTTTCAATTAATAAAGATTCATTTGATGATGATCAATAGTCCAGATTAC	103523
QY	538	ATATGCGATGTAACCAATCTGCGCAATATACAGGATCAAGCCGAGCATTAAGAGC	597
DB	103522	CAATGCGACGAACTATTAATTAAGAAATTAATCTTTTCACTATCAAAAAGAAAG	103463
QY	598	AAAGATGCGGCGGATTAAGCCCTTTGCTTTTGGCCAGGCTTAGGAACTGTGCTCTG	657
DB	103462	ATCGAACTATTAAGAAAGACATTAACCATTTAGTCAGAGTTCAAGTACTTTTGATTT	103403
QY	658	CCGGGCGATTAATACCGCGCTTCCGGTTTGTGACAGAGTGTATTATTAAGAAACATCTA	717

DB	103402	CCAGAGATTTTACACCTCTTCAAGATTGTTGAAACGGCTTATTAATAAACTACGCT	103343
QY	718	GAGCGCGCGCGCGATGAAGCAAGAGTGTAAACAGCGCTTTTCAATTTTGGCAATATG	777
DB	103342	GAGAAACCAAGTATTAATTAATGAGCGCTATTAGCGTTTGTCTATTAATTAATGATCTGTA	103283
QY	778	ACGATACCAAGGCGCGATGATTAAGCAAGAGCAAGAAATTCATTTATAGCAATATACT	837
DB	103282	ATTAATTCAAAAGAAATGTTATTAATGAGATGAGACAGCAAGTATTAATCACTTGTATTG	103223
QY	838	TCCGATGATGCAACGAACTGGAATCTATTTCCACCACTATGACATCGGAAATC	897
DB	103222	GCAATATATGCTCTCAAGAACTCTAATTAATTTTCACTTAATGTAATGAAGAAATC	103163
QY	898	CAAAAGTAATTAATTTATTCATGAAGACCTTGATGAGCTTAAAGTATTTCCGCT	957
DB	103162	AGAAATTAATTAATTTGTCGATTCCTTGAAATTAAGAAAGATTTAAACCTTCTATTT	103103
QY	958	AAAGCAGAGAGATTAATTCATGAGCTTAATTA	990
DB	103102	GTCATATGAGAAAGATTTTATGATTAATTA	103070

RESULT 11  
ABA90521\_19/c  
Continuation (20 of 24) of ABA90521 from base 190001 (Genomic sequence of Lactococcus  
WP Sequence split into 24 fragments LOCUS ABA90521 Accession ABA90521

WP	Fragment Name	Begin	End
WP	ABA90521_00	1	110000
WP	ABA90521_01	100001	210000
WP	ABA90521_02	200001	310000
WP	ABA90521_03	300001	410000
WP	ABA90521_04	400001	510000
WP	ABA90521_05	500001	610000
WP	ABA90521_06	600001	710000
WP	ABA90521_07	700001	810000
WP	ABA90521_08	800001	910000
WP	ABA90521_09	900001	1010000
WP	ABA90521_10	1000001	1110000
WP	ABA90521_11	1100001	1210000
WP	ABA90521_12	1200001	1310000
WP	ABA90521_13	1300001	1410000
WP	ABA90521_14	1400001	1510000
WP	ABA90521_15	1500001	1610000
WP	ABA90521_16	1600001	1710000
WP	ABA90521_17	1700001	1810000
WP	ABA90521_18	1800001	1910000
WP	ABA90521_19	1900001	2010000
WP	ABA90521_20	2000001	2110000
WP	ABA90521_21	2100001	2210000
WP	ABA90521_22	2200001	2310000
WP	ABA90521_23	2300001	2365589

Query Match 17.8% Score 177; DB 6; Length 110000;  
Best Local Similarity 50.8%; Pred. No. 3.2e-44;  
Matches 504; Conservative 0; Mismatches 480; Indels 9; Gaps 3;

QY	7	TCACAGCTTTACATTGGAACCTGCTACCGTAAACATGATTAGCAAGAACATGAT	66
DB	103402	TGTAGTATTGTTTACACTAGAGATCAAGATTAATTAACATTTCTTCTGTAACATGAT	103343
QY	67	TTTGATTTTCACTGGGACAGAGTGATCTCTATCCGCG---CCGTACAGCTGAAC	123
DB	103342	TTTCAATTAAGAAATGCGGACAAATTTCTTTATTCAGAAACAGAGATAGGATTT	103283
QY	124	AGTGAAGCTGACGGAAGGCGCCATCAGACAGTACGCGTTATGCGATGCG---GAGA	180
DB	103282	GCCCACTTAATGAAGAGACATTAACATCTTAATGCTTTTATGAGAGTGGGTGCGATG	3883
QY	181	AACTTGAATATATTTATTTGCCAGCGCATTAATGAAGCGGTTATCTTGCGGCG	240
DB	103102	GTCATATGAGAAAGATTTTATGATTAATTA	3823



QY 241 CTTTATTTTCGGGCTATGCGAGTACGAAAAACGATACGGAGATACCTTTCACATT 300  
 DB 3822 TTGTATTTTCCAGGCTATGCTGACTATGAGAAATTCAAAAAATTCAAAAGGAAATT 3763  
 QY 301 GTCCCGCATGATTTTGGATGATGGTCTGCTGACTCTGTGATCTTTTGGAAAGCTTAAA 360  
 DB 3762 TCACCTGATATGATTTTTCACAGGCTCTCACTCAACGAGTAAATTTAGAAATCATTT 3703  
 QY 361 GAAAAGATTCGATCTTTTAAAGATTTAGAAAAAATTAATCTA---TTGATACAGTT 417  
 DB 3702 GATTATTTGATATAAAAAATTTGTTATTTATTAATGATTAATCAACCTTAAAGTCTTACG 3643  
 QY 418 TTACGCTTCACTGATATTTGTACAGCCGAGCGCGGAAACCTGATGATGAACCAAGA 477  
 DB 3642 CCTCTTTCATTTTATTTTCTGATTTAGTGTCAAGTTGATTTGATTTGAACCAAG 3583  
 QY 478 GCGAGCGGCTCAAGCTATGATTAATCAACCTGTGTCTATGACGAAACGCGCCGACTTT 537  
 DB 3582 CAAGGTGACTTTCAATATTAATAAGATCAATGAGTATGACAAATAGTCCAGATTTAC 3523  
 QY 538 AATATGATATGATCAATCAATCTGACAAATATACAGATACGACGAAACATTAAGAGAG 597  
 DB 3522 CAATGACAGAAATCTAATTTAAGAAATTTCTTTTACCTTATCAAAAAGAAAGT 3463  
 QY 598 AAGAGATGCGCGGATTAAGCCCTTTCTGCTTTTGGCCAAAGCTTGAACCTGTGTCTG 657  
 DB 3462 ATCGAATCTTATGAGAAAGACATTAATAACATTTAGTCAAGTTCAGGTACTTTTGATTT 3403  
 QY 658 CCGGGGATTTATACACCGCTTCCCGCTTTGTACAGCTGTTTATTTGAAAGAACTCTA 717  
 DB 3402 CCAAGGATTTTACACCTCTTCAAGATTTGTTAGAACGCTTATTTAAAAAATCAACGCT 3343  
 QY 718 GAGCCGCGCGCGATGAAAGAAAGGTGTAACAGCGCTTTTCAATTTTGGCAATATG 777  
 DB 3342 GAGAAACCAAGTATTAATTAAGCCCTATTAAGCTTGTGATCATATTAATTAATCTGTA 3283  
 QY 778 ACGATACCAAGGCGCGAGTATACGAGAAAGAAATTCATTATACGCAATTAAT 837  
 DB 3282 ATTTATTTCCAAAAGAAATTTGTTATTAATCTGAGCATGAGCAAGTATTTCACTGTATTTG 3223  
 QY 838 TCCGATGATGACAGAACTGAAATCTAATTTCCACCACTATACATGCGCAATC 897  
 DB 3222 GATATATGATGCTCAGAAACTCTAAGTATTTATTTTCAACTTATGATATCAAGAAATC 3163  
 QY 898 CAATAAGTAAATTTATTTTATGATGAGACCTTGACTGTTTGAAGCCCTAAGTCTTCCGCT 957  
 DB 3162 AGGAAATTAATGTTTGTCAAGATTCCTTGAATAATGAGAAAGAAATTTAAAACTTTCTATT 3103  
 QY 958 AAGAGAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990  
 DB 3102 GTCAATGAGAGAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3070

## RESULT 12

ACF57507 standard; DNA, 1616 BP.

ACF57507;

22-APR-2004 (first entry)

B sphaericus penicillin-V-amidase DNA SEQ ID NO: 8.

Protein targeting; angiogenesis inhibitor; fusion gene; cancer; gene; ds;

antiangiogenic; cytostatic; ophthalmological; antiarthritic;

antiinflammatory; dermatological; immunosuppressive;

antiproliferative; antidiabetic; neuroprotective; vasotropic; anorectic.

Bacillus sphaericus.

MO2003093303-AL.

XX

XX

XX

XX

XX

13-NOV-2003.

06-MAY-2003; 2003MO-US014243.

06-MAY-2002; 2002US-0380063P.

(TEXA) UNIV TEXAS SYSTEM.

Hung M, Lan K, Ou-Yang F, Liu J, Lan K;

WPI; 2004-053038/05.

PT New polypeptide comprising an antiangiogenesis polypeptide region linked  
 PT to a therapeutic protein or a diagnostic protein, useful for diagnosing  
 PT or treating angiogenesis-dependent diseases e.g. cancer, lupus or  
 PT restenosis.

XX Disclosure; Page 87; Opp; English.

XX The present invention relates to a polypeptide comprising an  
 CC antiangiogenesis polypeptide region linked to a therapeutic protein or  
 CC polypeptide region, or a diagnostic protein or polypeptide region. The  
 CC polypeptide comprising an antiangiogenesis polypeptide region or the  
 CC nucleic acid encoding the polypeptide comprising the antiangiogenesis  
 CC polypeptide is useful for diagnosing or treating angiogenesis-dependent  
 CC diseases or cancer. Angiogenesis-dependent diseases include age-related  
 CC macular degeneration, arteriosclerosis, angiodysplasia, neovascular  
 CC glaucoma, arteriovenous malformations, nonunion fractures, arthritis,  
 CC rheumatoid arthritis, lupus, connective tissue disorders, Osler-Weber  
 CC syndrome, psoriasis, corneal graft neovascularisation, pyogenic  
 CC granuloma, delayed wound healing, retrolental fibroplasia, diabetic  
 CC retinopathy, scleroderma, granulations, haemangioma, trachoma,  
 CC haemophilic joints, vascular adhesions, hypertrophic scars, multiple  
 CC sclerosis, restenosis or obesity. The present sequence is a coding  
 CC sequence shown in the exemplification of the invention

SQ Sequence 1616 BP; 549 A; 244 C; 319 G; 504 T; 0 U; 0 Other;

Query Match 17.4%; Score 173.8; DB 12; Length 1616;

Best Local Similarity 49.8%; Pred No. 3.6e-44;

Matches 495; Conservative 0; Mismatches 492; Indels 6; Gaps 2;

QY 7 TGCACAGCTTACATGGAACCTGTCGACCGTAACATGATTAAGCAAGAAACATGAT 66  
 DB 335 TGCAGTAGCTATCAATTCGACACAGATGATTAAGTTATTCGCTGCAACATGAT 394  
 QY 67 TTTGATTTTACGCTGGGACAGAGTGTCTCTATCCGCGCGTTACAGCTGGAACAT 126  
 DB 395 TTTACAAATGGAACCAAGATGATTAAGTATTTGTCACACGTAATTAACGCAATTCGAT 454  
 QY 127 GAAAGTGAAGAA---GGGCCATCAGACAGATGACGGTTATGCGGTATGGGAGAA-- 181  
 DB 455 TTGAAAAAGAAAGATGATGATTAACAATTCATATGCTTTTGTGGAATGGAAGACAT 514  
 QY 182 -AACTTGAAATATATATTTTGCAGCGCAATTAAGAAAGCGTTATCTTTGCGGCG 240  
 DB 515 GACATTAATACACAGATCTCTATGATGGGGTAAAGAAAGGATTAATAGGCGCAATG 574  
 QY 241 CTTTATTTTCCGGGCTATGCGAGTACGAAAAACGATACGGAAGATACCTTACATT 300  
 DB 575 CTTTACTATGCTACATTTTGGCACTTATGCTGACGAACCTTAAAAAGGACAAACAGCATC 634  
 QY 301 GTCCGCAATGATTTTGTACATGAGTGTCTGCTGATGCTTCTTGGAAAGCTTAAA 360  
 DB 635 AATCCCGTATGATTAATTTTCAAGTTTATAGAAATTTGTAACTGTGATGATGATTTAT 694  
 QY 361 GAAAAGATTCATCTTTTACAGATTTGAGAAAGAAATTAAGATCTTATGATACATTTTA 420  
 DB 695 GAAAATTAATCTCTTATACATTTTGAATAGGCGCAATATTAATTAATCTTGGCTTGCACC 754  
 QY 421 CCGCTTCACTGATATTTGTACAGCCGAGCGGGCGAAACCTGACATAGAAACCAAGACA 480  
 DB 755 CCACTTCACTATTAATTTTACAGATGCTTGTGTGATGATGATTTTATTAACCGGATAAA 814



QY 481 GAGGCGCTCAAGTCTATGATATCAACCTGTGTCATGAGACAGCCCGACTTATA 540  
 DB 815 ACAGGCACTTACATTCATCGAAAACGATTGGCTCTATGAGAAATGCCCTGATAGA 874  
 QY 541 TGGCATGTAAACCAATCTGACGCAATATACAGAAATCAGACCGAATTAAGAGCAAA 600  
 DB 875 TGGCATGACAAATTTAAGAGCTTACATTTGGTCTACACCAAAATCCGCCACAGATATA 934  
 QY 601 GAGATGGGCGGATTAAGCCCTTCTGCTTTGGCCAGGCTTAGAAGCTGTGGCTGGCG 660  
 DB 935 ATGTGGGAGACTTGTGATTTGACACCGTTGGGCAAGGGGCGGCTTAGATTACCA 994  
 QY 661 GGGGATTATACACCGCTTCCCGGTTGTCTGACAGCTGTTTATTTGAAAACATCTAGAG 720  
 DB 995 GGTGATTTTACGCGCTGACGACGTTTCTTGGGTAGCATCTGAAAAAATATACGAA 1054  
 QY 721 CCGGCGGCGGATGAAACGAAAGGTGTAAACAGCCGCTTTCAAAATTTGGCAATATGACG 780  
 DB 1055 AAAGCCAAATGAAACAGAGGCGTAAACAACTTGTTCATTTCTATCTTCTGTAAT 1114  
 QY 781 ATACCAAAAGGCGGAGTATACGGAAGAAACGAAATTCATTTACCCATATATCTCC 840  
 DB 1115 ATCCCAAAAGGTGTGTTTGAACAAATGAGGGAACGAAATTAACCATCTATACCTCA 1174  
 QY 841 GTGATGTGCAACGAACTGAAACTATCTATTTCCACCACTATGACCAATCGGCAATCCAA 900  
 DB 1175 GCTATGTGTGCAACAAAGTAAACTATCTTAACTGTATGACAAATAGTGAATTTCA 1234  
 QY 901 AAAGTAAATTTATTTATCATGAAAGACCTTGACTGTGTTGAGCCTTAAAGTGTTCGCTAAA 960  
 DB 1235 GCCGTTTCCCTTAATGCTGAAATTTAATATGTCAGATTTAATTAATTTGAGTGGAT 1294  
 QY 961 GCAGAAAGACTATTCATGACTTAATTAAGA 993  
 DB 1295 CTTAAACAAAGTATTAAGCAATTAATCAAGTA 1327

RESULT 13  
 ID ADJ62865 standard; DNA; 1616 BP.

AC ADJ62865;  
 DT 06-MAY-2004 (first entry)  
 XX B sphaeriscus penicillin-V-amidase gene related to cancer treatment.  
 XX beta-catenin; bipartite T-cell factor; Tcf; promoter construct; Lef-1;  
 KW cytoskeletal; beta-catenin activity inhibitor; gene therapy; colon cancer;  
 KW metastasis; liver; thymidine kinase; prodrug; chemotherapy;  
 KW radiation therapy; surgery; penicillin-V-amidase; gene; ds.  
 XX Bacillus sphaeriscus.  
 OS US2003228285-A1.  
 FN 11-DEC-2003.  
 PD 05-MAY-2003; 2003US-00429802.  
 PF 03-MAY-2002; 2002US-0377672P.  
 PR (HUNG/) HUNG M.  
 PA (KMON/) KMONG K Y.  
 PA (ZOU/) ZOU Y.  
 XX Hung M, Kwong KY, Zou Y;  
 PI WPI, 2004-042209/04.  
 DR Novel viral vector comprising beta-catenin/bipartite T-cell factor-  
 PT responsive promoter having first and second promoter region linked to

PT target nucleic acid sequence, useful for treat- ing colon cancer.  
 XX disclosure; SEQ ID NO 17; 114pp; English.

PS This invention relates to a novel viral vector comprising a beta-  
 CC catenin/bipartite T-cell factor (Tcf)-responsive promoter construct which  
 CC contains a first promoter region having a copy of Tcf/Lef-1 binding site  
 CC operatively linked to a second promoter region, and a nucleic acid  
 CC sequence, where the first and second promoter regions are operatively  
 CC linked to target nucleic acid sequence. The invention may be useful for  
 CC the development of compounds with a cytoskeletal activity, through action  
 CC as beta-catenin activity inhibitors, or for gene therapy. The invention  
 CC may be useful for treating an individual with colon cancer which is  
 CC metastasized to the liver. The treatment involves administering the  
 CC vector of the invention where the nucleic acid sequence encodes a  
 CC therapeutic polypeptide or thymidine kinase, a prodrug and chemotherapy,  
 CC radiation, surgery or gene therapy to the individual. The present  
 CC sequence is that of a therapeutic gene which may be used in the vector of  
 CC the invention for the treatment of an individual with cancer.

XX Sequence 1616 BP; 549 A; 244 C; 319 G; 504 T; 0 U; 0 Other;

Query Match 17.4%; Score 173.8; DB 12; Length 1616;  
 Best Local Similarity 49.8%; Pred. No. 3.6e-44;  
 Matches 495; Conservative 0; Mismatches 492; Indels 6; Gaps 2;

QY 7 TGCACAGCTTCAATGTAAGTCTGACCGTAAACATGATATTAAGCAAGCAATGAT 66  
 DB 335 TGCAGTATGCTTATCAATTCGTACACAGATGATAAAGTTATTCGTCGCAATGAT 394  
 QY 67 TTTCATTTCACTGCGGACAGAGGTATCTCTATCCGGCGCTTACAGTGAACAGT 126  
 DB 395 TTTCATATGACACAGATGATTAAGTATTTGTCCAGTAAATTCAGGATTCGATTG 454  
 QY 127 GAGGTACAGGAA--GGGCCATCAGACACAGTACCGCTTATCGGTATGGGAGAA-- 181  
 DB 455 TTGAAAAAGAAATGATGCTTATACATTCATATGCTTTGTGGAATGGAAAGCACT 514  
 QY 182 -AAGCTGAAATATATATTTATTTGCGACGCGATTAATGAAAGCGTTATCTTGTGGCGG 240  
 DB 515 GACATTCATCAGACAGTCTCTATGATGGGTTAAACGAAAGGATTAATGGGCGCATG 574  
 QY 241 CTTTATTTTCCGGCTATGCGGAGTACGAAAAACGATACGGGAATCCGTTCACTT 300  
 DB 575 CTTTACTATGCTACATTTGCGACTTATGCTGACGAACTTAAAAAGCAACACGCGATC 634  
 QY 301 GTCCGATGAGTTTGTGACATGGGTGCTGCTGCTGCTGCTGCTTGTGGAAGCGTAAA 360  
 DB 635 AATCCCGTATGTAATTTCTCAAGTTTGAATTTGTGAACTGTGATGATGTTATT 694  
 QY 361 GAAAGATTCGATCTTTAAGCATTTGAGAAAAAATTAATGATCTATTGATACAGTTTA 420  
 DB 695 GAAAAATTAATCTTTTATATATTGATGAAAGCCAAATATTAATTAATGCTTGCACCC 754  
 QY 421 CCGCTTCACTGATATTGTACACCGGACGCGGCGGAAACCTGAGATGAAACCAAGACA 480  
 DB 755 CCACCTTCACTATATCAATTTACAGATGCTTGTGTAATCGATGTTATTAACCGGATAAA 814  
 QY 481 GAGGCGCTCAAGCTATGATATCAACCTGTGTCATGACGAAACAGCCCGCACTTTATA 540  
 DB 815 ACAGGCACTTACATTCATCGAAAACGATTGGCTCTATGAGAAATGCCCTGATAGA 874  
 QY 541 TGGCATGTAAACCAATCTGACGCAATATACAGAAATCAGACCGAATTAAGAGCAAA 600  
 DB 875 TGGCATGACAAATTTAAGAGCTTACATTTGGTCTACACCAAAATCCGCCACAGATATA 934  
 QY 601 GAGATGGGCGGATTAAGCCCTTCTGCTTTGGCCAGGCTTAGAAGCTGTGGCTGGCG 660  
 DB 935 ATGTGGGAGACTTGTGATTTGACACCGTTGGGCAAGGGGCGGCTTAGATTACCA 994  
 QY 661 GGGGATTATACACCGCTTCCCGGTTGTCTGACAGCTGTTTATTTGAAAACATCTAGAG 720  
 DB 995 GGTGATTTTACGCGCTGACGACGTTTCTTGGGTAGCATCTGAAAAAATATATCTGAA 1054

QY	721	CCGGGCGCGCATGAAACGAAAGGTGTAAACGGCGCTTTTCAAAATTTGGCAAAATATGAG	780
Db	1055	AAAGCGAAATATGAAACAGAAAGCGTTAACAACTGTTCATATTTCTTCTGTATAT	1114
QY	781	ATACCAAAAGGCGCGCATGTATACGAAAGACGAATTCATTATACGAATATATCTCC	840
Db	1115	ATCCCAAAAGGTGTGTGTTTGTGCAAAATGAGGGGAAACGGAATTATACATCATATCTCA	1174
QY	841	GTGATGTGTCAACGAATCTGAAACTACTATTTCCACGACTATGACCAATCGCAATCCGA	900
Db	1175	GCTATGTGTGCAACAAAGTAAATCTATTAACCTGTATGACAAATAGTCGAATTTCA	1234
QY	901	AAAGTAAATTTATTTTCATGAAAGCCCTGACTGTTTGAAGCCTAAAGTGTTTCCGCTAA	960
Db	1235	GCCGTTTCCTTAAGGTGTAATTTTAATATAGTCAAAGATTTAATATCATTTGAGTGGAAT	1294
QY	961	GCAGAGAGAGATATTCATGAGACTTAATTAAGA	993
Db	1295	CGTAAACAAGATATTTAGCAATTAATATCAAGTA	1327

AC	ACA46430	standard; DNA; 978 BP.
AC	ACA46430;	
DT	19-JUN-2003	(first entry)
DE	Prokaryotic essential gene #28087.	
KW	Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.	
OS	Staphylococcus epidermidis.	
PN	MO200277183-A2.	
XX	03-OCT-2002.	
PF	21-MAR-2002; 2002MO-US009107.	
PR	21-MAR-2001; 2001US-00815242.	
PR	06-SEP-2001; 2001US-00948993.	
PR	25-OCT-2001; 2001US-0342923P.	
PR	08-FEB-2002; 2002US-00072851.	
PR	06-MAR-2002; 2002US-0362659P.	
PA	(ELIT-) ELITRA PHARM INC.	
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;	
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;	
DR	WPI; 2003-029926/02.	
XX	P-PSDB; AB042560.	
PT	New antisense nucleic acids, useful for identifying proteins or screening	
PT	for homologous nucleic acids required for cellular proliferation to	
PT	isolate candidate molecules for rational drug discovery programs.	
XX	Claim 14; SEQ ID NO 34300; 1766pp; English.	
XX	The invention relates to an isolated nucleic acid comprising any one of	
XX	the 6213 antisense sequences given in the specification where expression	
XX	of the nucleic acid inhibits proliferation of a cell. Also included are:	
XX	(1) a vector comprising a promoter operably linked to the nucleic acid	
XX	encoding a polypeptide whose expression is inhibited by the antisense	
XX	nucleic acid; (2) a host cell containing the vector; (3) an isolated	
XX	polypeptide or its fragment whose expression is inhibited by the	
XX	antisense nucleic acid; (4) an antibody capable of specifically binding	
XX	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular	
XX	proliferation or the activity of a gene in an operon required for	

CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from MIP0 at  
CC ftp.wipo.int/pub/published\_pct\_sequences

Query Match	16.1%;	Score 160.2;	DB 8;	Length 978;
Best Local Similarity	49.9%;	Pred. 6.1e-40;		
Matches 455;	Conservative 0;	Mismatches 433;	Indels 6;	Gaps 2
QY	4	ATGTGCA	CAAGTCTTACATTGGAAAC	TGCTGACCGGTAAACATGATTTAGCAAGAACATG 63
DB	1	ATGTGTA	CGGCTATTTCTTATATATACAAA	CAACGTTATCATTTATTAGTACGAACGATG 60
QY	64	GATTTTGC	ATTTCAGCTGGGACAGAGGTATTTCTAT	CCGGCCGTTACAGCTGGAAC 123
DB	61	GACTTTGC	ATTTCGATTGAAATTTAATGTATCCCA	ACATGTTCCACCTCATTTACACATCCCA 120
QY	124	AGTGAAG	CTGACGGAAAGGCGCCATTCAGAC	CAGTACGCGTTTATCCGATATGGGAGAAA 183
DB	121	TTTGTATCT	AGATTCAC--	GACATGCGCTCTTGAATATGTTTTGTGGACAAATTTAAAA 177
QY	184	CTTGA	AAATATATTTATTTTGGCCGACCGCATTA	TGAAAGCGGTATTCCTGTGCGGCGCTT 243
DB	178	GTAGGA	CGTTATAGATTTTGGTGATGTGATTA	AGAAAAGSTTTAGCTATTTCCAAACAT 237
QY	244	TATTTTCC	GGCTATGCGGAGTACGAAAAACGATAC	GGGAAGATACGTTCAATGTC 303
DB	238	TACTTCA	CTGTGTAAGCGCTCATACAGTACCATTA	AAACGTTATGGTATTTTAACTTGAACA 297
QY	304	CCCGAT	GAGATTTTGTGACATGGGGTGTGCTG	CTGTCACTCTTTGGAAAGCTTAAAAGA 363
DB	298	CTGAGAG	AGTTATATGTTTGGGTTTTAGTTTTAA	TAAAGTATTTAGCGAATTTAAAAA 357
QY	364	AAGATT	CGATCTTTAA	CGATTTGAGTAAAGAAAATTTAGATCTTATTTGATACAGTTTACCG 423
DB	358	AAAGTT	AAAGATTCATATTTATGAATGA	AAAAAATTTACACTTTGAATATCGTTCCTT 417
QY	424	CTTCA	CTGATATTTGTCAACCGGACGGGCGAA	ACCTGACGATAGAACCAAGACAGAC 483
DB	418	TTTAC	ATTTTATGTGTCACTGATGA	AAACAGCATACGCTAGCATAGAACCTCAAAATGGC 477
QY	484	GGCCT	CAAAAGTCTATGATTAATCA	CACTGGTGTGATGACGAACACGCCGACTTTATATGG 543
DB	478	TTATTA	TATATGTATTAAGATATTAATTA	TGTTTACCTTTAACAAATGGCGCTTAATTTAGATTTGG 537
QY	544	CATGTA	CAAAATCTGACGAAATTA	CAGAAATCAGACCGAAGAAATTAGAGAGCAAAAG 603
DB	538	CATCTA	TCGATCTTAAAGAAATTA	TGTCTTATTTAACCCACAGAAATACCAATCAATTA 597
QY	604	ATAGG	CGGATTAAGCCCTTTCTGCTTTGGC	CAAGGCTTTAGGAATCTGTTGTCTGCCGGG 663
DB	598	ATAGG	TAAGTGTCTATGATTAAGATCA	TGGGTTGTGAAGCAGAAACAAATGGCTTACCGGGT 657
QY	664	GATTA	TACACCGGCTTCCGGGTTTTGTCA	AGGCTGTTTATTGAAAACATCTAGACCG 723

Db 658 GGTATATCTTCAACAGACGTTTATACGGCTACATTTTAAACACCACTACGCTGT 717  
Qy 724 GCGGCCCATGAAAGAAAGGTGTACACCCCTTTCAAAATTTGGCAAAATATACGATA 783  
Db 718 TCCCATATGAAAGTAAATTTATGAAATTTTAAAGTTCTAGATCAGTACGATC 777  
Qy 784 CCAAGGGCCGATGATTAACGAAAGACGAAATTCATTAATGCAATATACCTTCCGTG 843  
Db 778 CTTCAAGGTGAGATTAT---CGATGCCAATTAATTAACATTAACACATATCATTAAGT 834  
Qy 844 ATGTGCAACGAAGTGAAGTACTATTTCACCACTATGCAATTCGGCAATCCAAAA 903  
Db 835 ATGGAAGTAAAGAAAGCTATTATATTAAGCAGATCTTAGTAATCAATTTTCAA 894  
Qy 904 GTAAATTTATTCATGA 920  
Db 895 ATAAATTAATCTGAGA 911

## RESULT 15

AAH52510  
ID AAH52510 standard; DNA; 990 BP.

AAH52510;

03-SEP-2001 (first entry)

S. epidermidis open reading frame nucleotide sequence SEQ ID NO:413.

Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
endocarditis; de.

Staphylococcus epidermidis.

MO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000WO-US030782.

09-NOV-1999; 99US-0164258P.

(GLAXO) GLAXO GROUP LTD.

Kimmerly WJ;

WPI; 2001-316495/33.

P-PSDB; AAG81660.

Nucleic acid encoding polypeptides from Staphylococcus epidermidis,  
useful for vaccinating against infections, e.g. endocarditis.

Claim 8; Page 151-152; 2188pp; English.

AAH52304 to AAH5970 represent nucleic acids (I) encoding polypeptides  
(II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)  
and (II) can have antibacterial activity and therefore can be used in  
vaccination. The nucleic acids (I) may be used to produce the S.  
epidermidis polypeptides (II) via the production of vectors containing  
them which are used to produce host cells which express the  
polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
used to vaccinate subjects and to raise antibodies against the bacteria.  
The polypeptides may also be used to assay for other inhibitors of their  
activity and therefore identify compounds that may be used for the  
treatment of S. epidermidis infections, e.g. endocarditis. AAH5971 to  
AAH55090 represent specifically claimed S. epidermidis genomic DNA  
polynucleotide sequences from the present invention. AAH5091 to AAH5098  
represent oligonucleotide sequences and primers which are used in the  
exemplification of the present invention. N.B. The present invention  
specifically claims all the polynucleotide sequences given in the  
sequence listing of the present specification, however the sequence  
listing only goes up to SEQ ID NO:4454 so even though sequences are given

CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464  
XX  
SO Sequence 990 BP; 361 A; 158 C; 158 G; 313 T; 0 U; 0 Other;  
Query Match 16.0%; Score 159; DB 4; Length 990;  
Best Local Similarity 49.8%; Pred. No. 1,5e-39;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;  
Qy 2 ATATGCGCAAGATCTTACATGGAAGACCTGACCGCTTAACATGATTAAGCAAGACAA 61  
Db 8 ATATGCTATGCGCATTTCTTATATACAAACACGTTTCCATATTATAGTGAACAA 67  
Qy 62 TGAATTTGCAATTTCAAGCTGGGACAGAGGTGATTCATATCCGCGCTTACAGCTGA 121  
Db 68 TGGACTTTGCAATTTGATTTAATGATATCCCAACATGTTTCCACGCGCATATACATACC 127  
Qy 122 ACAATGAACTGACGGAAGGCGCCATACACAGATACGCTTTATGCGTAAGGAGAA 181  
Db 128 AATTTGATCTAGATTCAGACATCGCTCTT---GAATATGATTTTGTGGAACAAATTTAA 184  
Qy 182 AACTGGAATATATATATTTTCCGACGCGCATTAATGAAAGCGTTTATCTTGCGCGC 241  
Db 185 AAGTAGAGCTATATAGATTTGGATGATGTATTAACGAAAGGTTTACGTTATTCGAAC 244  
Qy 242 TTTATTTCCGCGCTATGCGAGTACGAAACAAAGATACGGAAGATACCGTTACATTTG 301  
Db 245 ATTAATTTCACTGATGGAAGCTCATACAGTACCAATAAAGTTATGTTATTTTAACTTAG 304  
Qy 302 TCCGCAATGATTTGACATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361  
Db 305 CACCTGAGAGTTATATGTTTGGGTTTATGTTTAAATTAAGATTAAGCAATTTAAAC 364  
Qy 362 AAAAGATGATCTTTAAGATGTGAGAAACAAATATAGATCTATGATACATTTTAC 421  
Db 365 AAAAGTTAAGAAAT 424  
Qy 422 CGCTTCACTGATATGTCAGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 481  
Db 425 CTTTACATTTCAATGTCAGTGAAGCAAGACGATACCGTATGAGACCTGACATG 484  
Qy 482 ACGGCTTCAAGTCTATGATATCAACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541  
Db 485 GCTTATTAATGATTAAGATATATATATATATATATATATATATATATATATATAT 544  
Qy 542 GGCATGTAACCAATCTGACAGATATACAGAAATCAGACGGAACATTAAGAGCAAG 601  
Db 545 GGCATCTATCTTAATTAAGAAATTAAGCTTTTAAAGCCACAGAAATCAACCAATCAAT 604  
Qy 602 AGATGCGCGATTAAGCCCTTTGCTTTTGGCCAAAGCTTAAGAACTGTGCTGCGG 661  
Db 605 TAAATAGTAAAGTCTGATGAAGATCATATGCGCTGTGAAGCAAGAACTGCTTACCG 664  
Qy 662 GGGATTAACACGCGCTTCCCGGTTTGTCAAGCTGTTTATTTGAAAGAACTTACAGC 721  
Db 665 GTGCTTATAGCTCAACAGATGTTTATTAAGCGCTTACATTTTAAGACACCACTAGCGT 724  
Qy 722 CCGGCGCGCAATGAAGAAAGGTGTAACAGCGCTTTCAAAATTTGGCAAAATTAAGCA 781  
Db 725 GTTCCATTAATGAAGATGAAGAAATTAATGAATTTTAAAGTTCTAGAAATCAAGTCA 784  
Qy 782 TACCAAGGCGCGAGTATACGGAAGACGAAATTCATTAATGCAATATACCTTCCG 841  
Db 785 TCCCTCAAGGTGAGATTAT---CGATGCCAATTAATTAATTAATTAATTAATTAAT 841  
Qy 842 TGAATGCAACGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 901  
Db 842 TGAATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 901  
Qy 902 AAGTAAATTTATTTTCATGA 920  
Db 902 AATTAATTAATTAATTAATGA 920

Mon Mar 20 10:09:57 2006

us-10-812-387-1.rng

Page 15

Search completed: March 18, 2006, 12:55:34  
Job time : 654 secs

---

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:46:51 ; Search time 4766 Seconds  
(without alignments)  
9777.577 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996  
Sequence: 1 catatgtgcacaagctctac.....atgagcttaataagaatcc 996

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	85.8	8.6	595	9 BH861030	BH861030 297 Lste
C 2	45.2	4.5	853	10 CZ540175	CZ540175 SRA-aad3
C 3	44.6	4.5	175	8 CV973894	CV973894 DPF370 Bt
C 4	44.6	4.5	881	9 AZ533653	AZ533653 ENTCS21R
C 5	44.6	4.5	910	9 AZ538681	AZ538681 ENTFO61R
C 6	44.6	4.5	923	9 BH155642	BH155642 ENTSS56TR
C 7	41.4	4.2	842	9 AZ688508	AZ688508 ENTIL247TR
C 8	41.4	4.2	871	9 AZ683787	AZ683787 ENTUK64TR
C 9	39.6	4.0	773	10 CZ859348	CZ859348 OC_Ba025
C 10	39.4	4.0	1055	9 CC254220	CC254220 CH261-4N2
C 11	38.6	3.9	512	9 BH059630	BH059630 RPCI-24-3
C 12	38.6	3.9	1101	10 CR008507	CR008507 Drosophila
C 13	38.2	3.8	575	11 CR008507	CR008507 Forward s
C 14	38.2	3.8	593	3 BQ266623	BQ266623 NISC FF14
C 15	38.2	3.8	853	9 AZ669903	AZ669903 ENTPI01R
C 16	38.2	3.8	859	9 AZ691994	AZ691994 ENTUN631R
C 17	38.2	3.8	881	9 AZ546452	AZ546452 ENTIP78TR
C 18	38.2	3.8	918	9 AZ688647	AZ688647 ENTIO56TR
C 19	38.2	3.8	929	9 BH149784	BH149784 ENTQD08TR
C 20	38	3.8	286	2 BH138826	BH138826 1029P41P
C 21	38	3.8	534	9 BH745273	BH745273 gU78C09.b
C 22	38	3.8	632	2 BB61316	BB61316 BB61316

23	38	3.8	678	4 AK085491	AK085491 Mus muscu
C 24	38	3.8	809	9 BH576107	BH576107 BOHNN15TR
C 25	38	3.8	997	5 BX838785	BX838785 BX38785
C 26	37.8	3.8	700	9 BH969780	BH969780 odc08606
C 27	37.6	3.8	625	9 BZ348527	BZ348527 hps8e09.g
C 28	37.6	3.8	670	9 CE125413	CE125413 C197-988
C 29	37.4	3.8	674	9 CC535077	CC535077 C197-988
C 30	37.2	3.7	953	10 CL089852	CL089852 ISB1-1481
C 31	37	3.7	439	7 CV163210	CV163210 rsmex1.00
C 32	37	3.7	403	1 AA190615	AA190615 zq44d03.r
C 33	37	3.7	606	10 CL326680	CL326680 RPCI44.24
C 34	37	3.7	713	9 BH97867	BH97867 oeg87d0I.
C 35	37	3.7	853	10 AG419422	AG419422 Mus muscu
C 36	36.8	3.7	197	11 CR829657	CR829657 GR0AA61C
C 37	36.8	3.7	427	1 AM829989	AM829989 ra48f11.y
C 38	36.8	3.7	435	1 AM829982	AM829982 ra48f03.y
C 39	36.8	3.7	482	1 AW735562	AW735562 ra21g07.y
C 40	36.8	3.7	505	1 AW783263	AW783263 ra26f02.y
C 41	36.8	3.7	505	1 AW783425	AW783425 ra28f01.y
C 42	36.8	3.7	505	1 AM828979	AM828979 ra70d09.y
C 43	36.8	3.7	505	1 AM828980	AM828980 ra70d10.y
C 44	36.8	3.7	505	1 AM828992	AM828992 ra70e11.y
C 45	36.8	3.7	505	1 AM829000	AM829000 ra70f07.y

## ALIGNMENTS

RESULT 1  
LOCUS BH861030/c 595 bp DNA linear GSS 12-FEB-2003  
DEFINITION 297 Listeria monocytogenes PUC18 library Listeria monocytogenes genomic clone 10-94 similar to Penicillin acylase/contagated bile hydrolyase and glutamate decarboxylase (adjacent Listeria monocytogenes ORFs), genomic survey sequence.

ACCESSION BH861030 GI:28317960  
VERSION BH861030.1  
KEYWORDS Listeria monocytogenes  
SOURCE Listeria monocytogenes  
ORGANISM Listeria monocytogenes  
REFERENCE 1 (bases 1 to 595)  
AUTHORS Call,D.R., Borucki,M.K. and Besser,T.E.  
TITLE Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of Listeria monocytogenes  
JOURNAL J. Clin. Microbiol. 41 (2), 632-639 (2003)  
PUBMED 12574259  
COMMENT Contact: Monica Borucki  
Animal Disease Research Unit  
USDA-ARS  
3003 ADRF, WSU, Pullman, WA 99164, USA  
Tel: 509 335 7407  
Fax: 509 335 8328  
Email: mborucki@vetmed.wsu.edu  
PCR PRIMERS  
FORWARD: M13F  
BACKWARD: M13R  
Insert Length: 595 Std Error: 0.00  
Seq primer: M13F, M13R  
Class: plasmid ends.  
Location/Qualifiers  
1..595  
/organism="Listeria monocytogenes"  
/mol\_type="genomic DNA"  
/strain="multiple"  
/db\_xref="taxon:1639"  
/clone="10-94"  
/clone\_1ib="Listeria monocytogenes PUC18 library"  
/note="Vector: pUC18; Shotgun library prepared by Amplicon Express (Pullman, WA)"

ORIGIN

Query Match 8.6%; Score 85.8; DB 9; Length 595;  
 Best Local Similarity 54.6%; Pred. No. 6,6e-14;  
 Matches 171; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 626 CTTTGGCCAGGCTTGAAGACTGTGTCGCCGGGAGATTATACACCGCTCCCGGT 685  
 |||||  
 DB 590 CATTTTCGACGATGACAGAAACAAATTAACAGGCTGTTATATCCGCCACAAAGCTT 531  
 |||||  
 QY 686 TTGTCAAGCTGTTATTTGAAAGAACTAGAGCCGCCGCCGATGAAACGAAAGTG 745  
 |||||  
 DB 530 TTGCTCCGTGGCGCATCTTGAAAGAAATATTAATTAAGCAAAACGAAAGAACCTA 471  
 |||||  
 QY 746 TAAACGCGCTTTCAATTTGGCAATATACGATACCAAGGCGCAGTGAATACG 805  
 |||||  
 DB 470 TCACAAATGTGTGTATGTCTTGAACAGTGTCCGATTCGAAATGAGCGGTATTAAG 411  
 |||||  
 QY 806 AAGAAAGCAAAATTCATTATACGAAATATACCTTCGTATGTGCAACGAACTGAAACT 865  
 |||||  
 DB 410 AAGTGGTGAACCTGATTTACACAAATGTGGCATCTATGTTCGGAAGCAAAAGCT 351  
 |||||  
 QY 866 ACTTTTCCACACTATGACAAATCGCAATCCAAAGTAAATTTATTCATGAAGAC 925  
 |||||  
 DB 350 ACTATTTCACTCATACGAAACCAACCAATCAATAGTTACTTTGCGAAGAAAGTTA 291  
 |||||  
 QY 926 TTGACTGTTTGA 938  
 |||||  
 DB 290 TAGAAATATGAA 278

RESULT 2  
 C2540175/c 853 bp DNA linear GSS 13-MAY-2005  
 LOCUS  
 DEFINITION SRA-aad30b09.b1 Strongyloides ratii whole genome shotgun library  
 (SRASS 004) Strongyloides ratii genomic, genomic survey sequence.

ACCESSION C2540175  
 VERSION C2540175.1 GI:64664074  
 KEYWORDS GSS.

SOURCE Strongyloides ratii  
 ORGANISM Strongyloides ratii  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
 Panagrolaimoidea; Strongyloidea; Strongyloides.

REFERENCE 1 (bases 1 to 853)  
 Miteva,M., McCarter,J.P., Thompson,F., Viney,M., Pape,D.,  
 Rafter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,  
 Clifton,S.W. and Wilson,R.

TITLE Genome Survey sequences from the rat parasitic nematode  
 Strongyloides ratii  
 JOURNAL Unpublished (2005)  
 COMMENT Contact: Miteva M

Washington University in St. Louis  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: nematode@wustl.edu  
 Genomic DNA was provided by Fiona Thompson  
 (F.Thompson@bristol.ac.uk) and Mark Viney  
 (Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,  
 UK.

Class: shotgun.  
 Location/Qualifiers  
 1..853

FEATURES  
 source  
 /organism="Strongyloides ratii"  
 /mol\_type="genomic DNA"  
 /strain="isofemale line ED321 heterogenic"  
 /db\_xref="taxon:34506"  
 /dev\_stage="infective larval stage (L3)"  
 /lab\_host="GSI0"  
 /clone\_lib="Strongyloides ratii whole genome shotgun  
 library (SRASS 004)"  
 /note="Vector: pOTW13; Site\_1: BscXI; Site\_2: BscXI;  
 Strongyloides ratii genomic DNA was randomly sheared,  
 end-repaired and size fractionated to enrich for 2-4 kb

fragments. Genomic DNA was provided by Fiona Thompson  
 (F.Thompson@bristol.ac.uk) and Mark Viney  
 (Mark.Viney@bristol.ac.uk) at the University of Bristol,  
 Bristol, UK. Sequencing by Washington University Genome  
 Sequencing Center, St. Louis, MO."

ORIGIN  
 Query Match 4.5%; Score 45.2; DB 10; Length 853;  
 Best Local Similarity 54.9%; Pred. No. 0.098;  
 Matches 89; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 408 GATACAGTTTACCGCTTCACTGATATTTGCAGACCGGACGGCCGAACCTGACAT 467  
 |||||  
 DB 257 GAATTAACATCACTCATCTATGCTTTTATGATTAACACGGTGTGCAATTATGT 198  
 |||||  
 QY 468 AAGAACAGAGCAGACGGCTCAAGTCAATGATTAATCAACCTGTGTGACGAAACAG 527  
 |||||  
 DB 197 TGAAGCGTAAATGCAAAATTACAGGTTTGTAAACCTCACTCGTGTCACTAATGG 138  
 |||||  
 QY 528 CCCCAGCTTATATGCGATGATTAACCAATCTGCAGCAATATAC 569  
 |||||  
 DB 137 GCCTGATTTCAATGCAATTTAATAAACCTGAATTAATATAC 96

RESULT 3  
 CV973894 175 bp mRNA linear EST 01-JAN-2005  
 LOCUS DDF370 Brassica oleracea DD-PCR fragment Brassica oleracea cDNA 3',  
 mRNA sequence.

ACCESSION CV973894  
 VERSION CV973894.1 GI:56957135  
 KEYWORDS EST.

SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 175)  
 Casimiro,S., Ze-Ze,L., Tenreiro,R. and Monteiro,A.A.  
 Pathogenesis-related cDNA genes isolated by differential-display  
 PCR from Brassica oleracea seedlings infected with Peronospora  
 parasitica

JOURNAL Unpublished (2003)  
 COMMENT Contact: Casimiro, S.  
 Departamento de Biologia Vegetal - Centro de Genetica e Biologia  
 Molecular  
 Faculdade de Ciencias da Universidade de Lisboa  
 Rua Ernesto de Vasconcelos, Rd. C2, Piso 4, 1749-016 Lisboa,  
 Portugal

Fax: +351217500048  
 Email: scasimiro@fc.ul.pt  
 Insert Length: 175 Std Error: 0.00  
 Seq primer: T3  
 High quality sequence stop: 175  
 POLYA=yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..175

/organism="Brassica oleracea"  
 /mol\_type="mRNA"  
 /cultiivar="CRG3.1", 'Algarvia' and 'Coracao-de-Boi'  
 /db\_xref="taxon:3712"  
 /dev\_stage="seven days after infection with Peronospora  
 parasitica"  
 /clone\_lib="Brassica oleracea DD-PCR fragment"

ORIGIN  
 Query Match 4.5%; Score 44.6; DB 8; Length 175;  
 Best Local Similarity 61.7%; Pred. No. 0.092;  
 Matches 71; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 467 TAGAACAGAGCAGACGGCTCAAGTATGATTAATCAACCTGTGTGACGAAACA 526  
 |||||  
 DB 132 TCGAACCAACGAAATGAATCAAAATTTATGATCACTCAATCGAGTATGCAATA 73

QY 527 GCCCCACTTTATATGCGATGTACCAATCTGCAGCAATATACAGCAATCACC 581  
DB 72 GTCTGATTATCTCTGGCATGAAATAATTTAGGAATATCTTTCAATCAATCACC 18

RESULT 4  
A2533653/c 881 bp DNA linear GSS 03-NOV-2000  
LOCUS ENTSC32TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION genomic, genomic survey sequence.  
ACCESSION A2533653  
VERSION A2533653.1 GI:11089512  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
REFERENCE 1 (bases 1 to 881)  
Loftus,B., Van Aken,S. and Fraser,C.  
Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library  
Unpublished (2000)  
JOURNAL Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: b.loftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Classes: shotgun  
High quality sequence start: 21  
High quality sequence stop: 842.

FEATURES  
source

1..881  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHD1; Site 1: Bat I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaubin and B.  
Bartell, Oxford University Press, 1999)."

ORIGIN

Query Match 4.5%; Score 44.6; DB 9; Length 881;  
Best Local Similarity 49.8%; Pred. No. 0.15;  
Matches 113; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
QY 383 TTGTAGAGAAAAATTTAGATCTATTTGATACAGTTTACCGCTTCAGTGAATTTGTCAG 442  
DB 632 TTGTATAGCAAAATATCAACAGTAGAATATCTTACTATCTTGTATTTATCATTAATA 573  
QY 443 ACCGACGGCGCGGAACCTGACATAGAACCAAGACGACGCGCTCAAGTCTATGATA 502  
DB 572 ATCACAATATGTAAGAAACATGACATATTTGGCATATCATATGAGGAATTAATAAGAA 513  
QY 503 ATCAACTGTGTGATGACGAACAGCCCGACTTTATATGCGATGTACCAATCTGCAGC 562  
DB 512 ATACACTGGAATATATCATATTAATAAAACCCCATATGGAACAATGTTCCATATTAAGA 453

QY 563 AATATACAGGAATCAGACCGCAATTAAGAGCAAGATGGC 609  
DB 452 CATTACAAATGATATTGGAACATACAAATATATGATTAAGATGATGAC 406

RESULT 5  
A2538681/c 910 bp DNA linear GSS 14-NOV-2000  
LOCUS ENTSC06TRF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION genomic, genomic survey sequence.  
ACCESSION A2538681  
VERSION A2538681.1 GI:11143851  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
REFERENCE 1 (bases 1 to 910)  
Loftus,B., Van Aken,S. and Fraser,C.  
Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library  
Unpublished (2000)  
JOURNAL Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: b.loftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Forward  
Classes: shotgun  
High quality sequence start: 50  
High quality sequence stop: 716.

FEATURES  
source

1..910  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHD1; Site 1: Bat I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaubin and B.  
Bartell, Oxford University Press, 1999)."

ORIGIN

Query Match 4.5%; Score 44.6; DB 9; Length 910;  
Best Local Similarity 49.8%; Pred. No. 0.15;  
Matches 113; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
QY 383 TTGTAGAGAAAAATTTAGATCTATTTGATACAGTTTACCGCTTCAGTGAATTTGTCAG 442  
DB 370 TTGTATAGCAAAATATCAACAGTAGAATATCTTACTATCTTGTATTTATCATTAATA 311  
QY 443 ACCGACGGCGCGGAACCTGACATAGAACCAAGACGACGCGCTCAAGTCTATGATA 502  
DB 310 ATCACAATATGTAAGAAACATGACATATTTGGCATATCATATGAGGAATTAATAAGAA 251  
QY 503 ATCAACTGTGTGATGACGAACAGCCCGACTTTATATGCGATGTACCAATCTGCAGC 562  
DB 250 ATACACTGGAATATATCATATTAATAAAACCCCATATGGAACAATGTTCCATATTAAGA 191  
QY 563 AATATACAGGAATCAGACCGCAATTAAGAGCAAGATGGC 609





RESULT 8  
LOCUS A2683787/c 871 bp DNA linear GSS 14-DEC-2000  
DEFINITION ENTX64TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
ACCESSION A2683787  
VERSION A2683787 GI:11820933  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
REFERENCE 1 (bases 1 to 871)  
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
JOURNAL HMI:IMSS sheared DNA library  
COMMENT Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: b.loftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 25  
High quality sequence stop: 860.  
Location/Qualifiers  
1..871  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: PHOS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

## FEATURES

source

## ORIGIN

Query Match 4.2%; Score 41.4; DB 9; Length 871;  
Best Local Similarity 48.9%; Pred. No. 1.3; Mismatches 0; Gaps 0;  
Matches 111; Conservative 0; Indels 116; Indels 0; Gaps 0;

QY 383 TTGTAGAGAAAATTAGATCTATGTGATACAGTTTACCGCTTCACTGGATTTGTCG 442  
DB 500 TTGTAAAGACAAATATACACAGTAGACATTTCTATCTTGTATGTCATAAAA 441  
QY 443 ACCGAGCGGGCCGAAACCTGACATAGAACCAAGAGAGAGCGCTCAAGTCTATGATA 502  
DB 440 ATCACAATATGAAAAACATGACCATATAGCATTTATATATGGAGGATTAATAGAAA 381  
QY 503 ATCAACCTGTGTCATGACGACAGCCCGCATTTATATGCGCATGTAACTCAATCTGACG 562  
DB 380 ATATACAGAAATATATCAATATTAATAAACCATATGAAACAGTGTCCAAATATAAGA 321  
QY 563 AATATACAGGAATCAGACCGAGCAATTTAGAGAGCAAGAGATGGCG 609  
DB 320 CATCAATATGATTTTAAACAAATATATGCTATTAAGTATGATGAC 274

RESULT 9  
LOCUS C2859348/c 773 bp DNA linear GSS 27-JUL-2005  
DEFINITION OC\_Ba0251M04.f OC\_Ba Oryza coarctata genomic clone OC\_Ba0251M04  
ACCESSION C2859348  
VERSION C2859348 GI:71334104  
KEYWORDS GSS.  
SOURCE Oryza coarctata (Porteresia coarctata)  
ORGANISM Oryza coarctata  
REFERENCE 1 (bases 1 to 773)  
AUTHORS Kim, H., Collura, K., Wisniewski, M., Byrne, M., Stum, D., Smart, D.,  
Rao, K., Luo, M., Uetzy, R., Kudrna, D., Muller, C., Soderlund, C. and  
Wing, R.  
TITLE Oryza Map Alignment Project - Arizona Genomics Institute  
JOURNAL Unpublished (2005)  
COMMENT Contact: Rod A. Wing  
Arizona Genomics Institute  
University of Arizona  
Forbes Building Room 303, Tucson, AZ 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: rwing@genome.arizona.edu  
PCR PRIMERS  
FORWARD: TAA TAC GAC TCA CTA TNG GG  
BACKWARD: CAC TCA TTA GGC ACC CCA  
Plate: 0251 row: M column: 04  
Seq primer: TAA TAC GAC TCA CTA TNG GG  
Class: BAC ends.  
Location/Qualifiers  
1..773  
/organism="Oryza coarctata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:77588"  
/clone\_lib="OC\_Ba0251M04"  
/tissue\_type="leaves"  
/dev\_stage="mature"  
/lab\_host="DH10B"  
/clone\_lib="OC\_Ba"  
/note="Vector: pGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

## FEATURES

source

## ORIGIN

Query Match 4.0%; Score 39.6; DB 10; Length 773;  
Best Local Similarity 48.3%; Pred. No. 4.5; Mismatches 119; Indels 0; Gaps 0;  
Matches 111; Conservative 0; Indels 119; Indels 0; Gaps 0;

QY 716 TAGAGCGGCGCGCGATGAACGAAGGTGTACAGCCGCTTTCAATTTTGGCAATA 775  
DB 230 TAAACACGATTTGCTGTGCTGAGCTATCAACCGGGAGGTGTGAAGTGAACCCAAC 171  
QY 776 TGAAGTATCCAAAGGCGGAGTATACGGAAGAGACGAATTTCTATTATAGCAATATA 835  
DB 170 TACTGCAACACAAAGTAAATATACAGAGTAAAGTTGCACTACTTATTATGAAATTT 111  
QY 836 CTTCGCTGATGTGCAACGAAGAACTGAAATCTACTATTTCACCACTATGCAATCGGCA 895  
DB 110 GCACTATCATTTGCACTGCAAAAATATTAATTTTATATGCAATTTAGAGATGGCCAAA 51  
QY 896 TCCAAAAGTAATTTATTTATTTATGAAAGACTGACGTGTTGGAGCTTAA 945  
DB 50 TGAAGATTAATAAATCATGCTCTATATAGAGTTTCAATGCGCAAGATTA 1

RESULT 10  
LOCUS CC254220/c 1055 bp DNA linear GSS 13-MAY-2003  
DEFINITION CH261-4N24\_Spe.1 CH261 Gallus gallus genomic clone CH261-4N24,  
ACCESSION CC254220 genomic survey sequence.

VERSION CC254220.1 GI:30590970  
 KEYWORDS GSS  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 1055)  
 Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
 Warren, W., Graves, T., Maritz, E. and Wilson, R.  
 Gallus gallus BAC End Reads  
 Unpublished (2003)  
 Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert Length: 18200 Std Error: 0.00  
 Seq primer: SP6 ATTAGGTGACACTATAG  
 Class: BAC ends  
 High quality sequence start: 113  
 High quality sequence stop: 391.  
 Location/Qualifiers  
 1..1055  
 /organism="Gallus gallus"  
 /mol\_type="genomic DNA"  
 /strain="Red Jungle Fowl"  
 /db\_xref="taxon:9031"  
 /clone="CH261-4N24"  
 /sex="female"  
 /cell\_line="UCD001, inbred 256"  
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CH261 Female Chicken library - For library and clone  
 ordering information: http://www.chori.org/bacpac"

ORIGIN  
 Query Match 4.0%; Score 39.4; DB 9; Length 1055;  
 Best Local Similarity 48.4%; Pred. No. 5.6;  
 Matches 109; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 694 GCTGTTATTTGAAAGACATCTAGAGCGGCGCGGTAAGAAAGGTTACACGCC 753  
 DB 376 GCCTTCAGTAAAGGTGAGTACTAGACAGTGAACACAGACAGTGCACAAAAGT 317  
 QY 754 GCTTTCAATTTTGCAATATGACATACCAAGGCGGCAATGATACGGAAGAAC 813  
 DB 316 GCTGATGAAATCAAGGATCTTCTGTATACAAAAGAGTTATGTTAGGATTAAGCA 257  
 QY 814 GAATTTATTTATGACATATATCTCCGTGATGTGCAAGAACTGAAACTATATTTC 873  
 DB 256 TAAATAGCTAGTTCCTCAATTAAGTTGAAGATAGACCGTACGAAAATTAATATTT 197  
 QY 874 CACCATATGACATCGCAATCCAAAAGTAAATTTATTTTCAT 918  
 DB 196 CTTCTACACACAGCAGCGTACCTTCAGACATATTTTCAAT 152

RESULT 11  
 BH059630 512 bp DNA linear GSS 18-JUL-2001  
 LOCUS RPCI-24-381C20.TV RPCI-24 Mus musculus genomic clone  
 DEFINITION RPCI-24-381C20, genomic survey sequence.  
 ACCESSION BH059630  
 VERSION BH059630.1 GI:14869888  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Muridae; Murinae; Mus.  
 1 (bases 1 to 512)  
 Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akintun, B., Levins, M.,  
 Tseng, G., Geer, K., Kroll, M., Shvartbey, A., Gebregeorgis, B.,

TITLE Russell, D., de Jong, P. and Fraser, C.M.  
 JOURNAL Mouse BAC End Sequences from Library RPCI-24  
 COMMENT Unpublished (1999)  
 Other GSS: RPCI-24-381C20.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-24. For BAC  
 library availability, please contact Pieter de Jong  
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC  
 Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
 Plate: 381 row: C column: 20  
 Seq primer: T7  
 Class: BAC ends.  
 Location/Qualifiers  
 1..512  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-24-381C20"  
 /sex="male"  
 /cell\_line="Spleen/Brain"  
 /note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI.  
 RPCI-24 Mouse BAC library produced by Pieter de Jong. The  
 library was cloned in the pTARBAC1 cloning vector at the  
 BamHI sites using MboI partially digested male C57BL/6J  
 DNA."

ORIGIN  
 Query Match 3.9%; Score 38.6; DB 9; Length 512;  
 Best Local Similarity 53.7%; Pred. No. 7.8;  
 Matches 80; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 727 GCCATGAAGAAAGGTGTAAGCGCGCTTTCAATTTTGCAATATGACATACCA 786  
 DB 103 GCCAAATTAATTAAGGCGCCAGGAGGAGTATCCAAATTTAAATTAAGAAAGCA 162  
 QY 787 AAGGCGCAGTATTAAGCAAGAAAGCAATTTATTAATGCAATATCTCCGTATG 846  
 DB 163 AAGGAGTTATTAACAAGAAATGAGAAATTAACAAAATATATCAACAACATGAAAA 222  
 QY 847 TGCAACGAACCTGAAACTATATTTC 875  
 DB 223 TCTATGAATGAATGGAAGTTTCTATACA 251

RESULT 12  
 CNS00FOO 1101 bp DNA linear GSS 03-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:  
 DEFINITION BACR31P03 of RPCI-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 ACCESSION AL070854  
 VERSION AL070854.1 GI:4950896  
 KEYWORDS GSS.  
 SOURCE Drosophila melanogaster (fruit fly)  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1101)  
 Genoscope.  
 Direct Submission  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)



/note="Organ: mammary; Vector: PCMV-SPOrt6; Site 1: Sali;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hemmighausen/Robin Humphreys,  
NIH"

## ORIGIN

Query Match 3.8%; Score 38.2; DB 3; Length 593;  
Best Local Similarity 49.3%; Pred. No. 11;  
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 721 CCGGGGCGGATGAAACGAAAGTGTAAACGCGCTTTTCAATTTGGCAATATGACG 780  
Db 192 CCCCCCCCCCAAAAAAAAAAAAAAAAAACCCCTTTTAAAAAAAAAAAAAAAAACCTT 251  
Qy 781 ATACCAAGGGGCGGATGATTAACGGAAGACGAAATTCATTAATGCAATATATCTTC 840  
Db 252 TTTAAAGGGGCAAAAAAAAAATTTGGGGGAAAAAAAAAGGGGGGAAAACTTGG 311  
Qy 841 GTGATGTCACGAACTGAAACTACTATTTCCACCACTATGACATCGCAATCCAA 900  
Db 312 GAAAAAGAAAAAAATTTAAATAATTTTAAAAAAAAAAAAAAAAAGGCTTTTAA 371  
Qy 901 AAGTAAATTTATTTTCATGAAGA 923  
Db 372 AAAAAAATTTTAAAAAAAAA 394

RESULT 15  
AZ669903/c 853 bp DNA linear GSS 14-DEC-2000

LOCUS ENTHPIOTF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
DEFINITION genomic, genomic survey sequence.

ACCESSION AZ669903  
VERSION AZ669903.1 GI:11807049

KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 853)

AUTHORS Loftus, B., Van Aken, S. and Fraser, C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HMI:IMSS sheared DNA library

COMMENT Unpublished (2000)

CONTACT: Brendan J Loftus

Department of Eukaryotic Genomics

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: b.loftus@igr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Forward

Class: Shotgun

High quality sequence start: 66

High quality sequence stop: 746.

Location/Qualifiers

1. 853

/organism="Entamoeba histolytica"

/mol\_type="genomic DNA"

/strain="HMI:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: pHOsi, Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, I.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.) The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

## ORIGIN

Query Match 3.8%; Score 38.2; DB 9; Length 853;  
Best Local Similarity 48.0%; Pred. No. 12;  
Matches 109; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 383 TTGTAGAGAAAAATAGATCTATTTGATATACGCTTCACTGATATTTGTCAG 442  
Db 525 TTGTAAAGCAAAATATCAACATGAAACATTTCTATTTGTGTCATTAAGA 466  
Qy 443 ACCGAGCGGCGGAACTGACATGAAACCAAGAGAGACGCGCTCAAGTATGATA 502  
Db 465 ATCAAAATGAAAAACATGACCATATGAGCATATATATGAGAGGATTAAGA 406  
Qy 503 ATCAACTGTGTATGACGAAACGCGGACCTTATATGTCATGTAACCATGTCAGC 562  
Db 405 ATCACTGAAATTAACATATTAATAAACCCTCATATGAAACAGTGTTCATTAAGA 346  
Qy 563 AATATAGGAATGACCGAGCAATTAGAGCAAGAGATGAGC 609  
Db 345 CATCAATGATATGAAACAAATCATGATTAAGTATGAC 299

Search completed: March 18, 2006, 13:11:47  
Job time : 4770 secs

H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:46:58 ; Search time 226 Seconds

(without alignments)  
7833.856 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996  
Sequence: 1 catatgtgcacaaagctctac.....atgagcttaataagatcc 996

Scoring table: IDENTITY NUC  
Gapop 10-0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/6C.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/6D.COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/6E.COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/6F.COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/6G.COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	16.0	990	US-09-710-279-413	Sequence 413, App
2	159	16.0	996	US-09-134-001C-2421	Sequence 2421, Ap
3	159	16.0	2975	US-09-710-279-3345	Sequence 3345, Ap
4	159	16.0	4114	US-09-710-279-4184	Sequence 4184, Ap
5	146	14.7	11466	US-08-956-171E-444	Sequence 444, App
6	146	14.7	11466	US-08-781-986A-444	Sequence 444, App
7	95.4	9.6	987	US-09-104-500C-838	Sequence 838, App
8	82.6	8.3	978	US-09-137-32A-366	Sequence 366, App
9	47.8	4.8	400	US-08-956-171E-3720	Sequence 3720, Ap
10	47.8	4.8	400	US-08-781-986A-3720	Sequence 3720, Ap
11	39	3.9	7218	US-08-232-463-14	Sequence 14, Appl
12	37.4	3.8	107140	US-09-949-016-14834	Sequence 14834, A
13	36.4	3.7	61663	US-09-453-702B-62	Sequence 62, Appl
14	36.4	3.7	61663	US-10-114-170-62	Sequence 62, Appl
15	34.4	3.5	4429	US-09-455-486-7	Sequence 7, Appl
16	34.2	3.4	2067	US-09-106-194-11	Sequence 11, Appl
17	34	3.4	505	US-09-621-976-15639	Sequence 15639, A
18	34	3.4	1083	US-09-134-000C-2381	Sequence 2381, Ap
19	33.4	3.4	295	US-09-213-294A-5605	Sequence 5605, Ap
20	33.2	3.3	872	US-09-270-767-25560	Sequence 25560, A
21	33.2	3.3	2228	US-09-270-767-10203	Sequence 10203, A
22	33.2	3.3	2460	US-08-952-165-1	Sequence 1, Appl
23	33.2	3.3	6065	US-09-800-729-35	Sequence 35, Appl
24	33.2	3.3	101011	US-09-949-016-16933	Sequence 16933, A

C 25	33	3.3	1801	3	US-08-669-304-30	Sequence 30, Appl
C 26	33	3.3	1801	3	US-09-824-053-30	Sequence 30, Appl
C 27	33	3.3	247299	3	US-09-949-016-17590	Sequence 17590, A
C 28	32.6	3.3	1200	3	US-09-107-532A-19	Sequence 19, Appl
C 29	32.6	3.3	1509	3	US-09-248-796A-3224	Sequence 3224, Ap
C 30	32.4	3.3	601	3	US-09-949-016-113423	Sequence 113423, A
C 31	32.4	3.3	601	3	US-09-949-016-163399	Sequence 163399, A
C 32	32.4	3.3	601	3	US-09-949-016-163400	Sequence 163400, A
C 33	32.4	3.3	601	3	US-09-949-016-163401	Sequence 163401, A
C 34	32.4	3.3	1340	3	US-08-956-171E-245	Sequence 245, App
C 35	32.4	3.3	1340	3	US-08-781-986A-245	Sequence 245, App
C 36	32.4	3.3	85869	3	US-09-949-016-12017	Sequence 12017, A
C 37	32.4	3.3	85878	3	US-09-949-016-16321	Sequence 16321, A
C 38	32.4	3.3	167708	3	US-09-949-016-16423	Sequence 16423, A
C 39	32.4	3.3	285966	3	US-09-949-016-12287	Sequence 12287, A
C 40	32.4	3.3	288031	3	US-09-949-016-14864	Sequence 14864, A
C 41	32.2	3.2	832	3	US-09-621-976-2813	Sequence 2813, Ap
C 42	32.2	3.2	255679	3	US-09-949-016-17189	Sequence 17189, A
C 43	32.2	3.2	786431	3	US-09-751-389-3	Sequence 3, Appl
C 44	32	3.2	870	3	US-09-543-681A-3984	Sequence 3984, Ap
C 45	32	3.2	930	4	US-09-605-703B-351	Sequence 351, App

## ALIGNMENTS

RESULT 1  
US-09-710-279-413  
Sequence 413, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PUA480US  
CURRENT APPLICATION NUMBER: US/09/710, 279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164, 258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 413  
LENGTH: 990  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-413

Query Match 16.0%; Score 159; DB 3; Length 990;  
Best Local Similarity 49.8%; Pred. No. 1.4e-39;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;  
QY 2 ATATGTGACAAAGCTTACATTTGAAAGCTGACCGTAACATGATTAGCAAGACAA 61  
8 ATATGTGACGACGATTTCTTATATATACAAACAGTTACATTTATTAGTAGAACA 67  
DB 62 TGAATTTTCATTTGACGCTGGGACAGAGCTATTCTATCCGCCCGTTACAGTGA 121  
68 TGAATTTTCATTTGATTAATTAATGATCCCAACATTTTCACGCCCATTTACCTACC 127  
QY 122 ACAGTGAACCTACAGGAGGCCCATACAGACAGTACCGTTATTCGATAGGGAGAA 181  
128 AATTGATCTAATTCAGACATGCGCTT---GAATATGTTTGTGGAACAATTAA 184  
QY 182 AACTTGAAATATATTTATTTGCCGACGCAATTAATGAACGCGTTATCTTGGCGGC 241  
185 AAGTAGAGATGATATGATTTGATGATGATTAACCAAAAGGTTAGCTATTTGCAAC 244  
QY 242 TTTATTTTCGCGGCTATGCGAGTACGAAACATATAGGGAAGATACGTTACATG 301  
245 ATTACTTACTGATGAGAGCTTACATACAGTACCATTAACGTTATTTAACTTAG 304

Oy	302	TCGCCATGAGTTTGGACATGGGCTCTCAGTCTGCAGCTCTTGGAAACGTAAG	361
Oy	303	CACTGAGGATTATTTGTTGGGTTTGGTTTAAATGAATATGCGAATTTAAAC	364
Oy	362	AAAAGATTCGATCTTTAATCGATTGTAGAGAAAAATTAAGATCTATTTGGATACAGTTTAC	421
Db	365	AAAAGGTTAAGAAATCAATATTATGATGAAATGAAAAAATATACGCTTTGAATATCGTCTC	424
Oy	422	CGCTTCACTGGAATTTGTCAAGCCGACGGGCCGAAACTGACGATAGAACCAAGACAG	481
Db	425	CTTACATTTTCATGGTCACTGATGAAGAAACAGACATACCGTACCCTAATACCTCAATG	484
Oy	482	ACGGCCCAAGTCTATGATTAATCAACCTGGTGTATGACGAAACAGCCCGACCTTATAT	541
Db	485	GCTTATTTAATAGTTTAAAGATTAATGTTTCATACCTTAACAAATGAACCTTAATTAAGTT	544
Oy	542	GGCATGTAAACCAATCTGCAGCAATATACGAATACAGCCGAAGCAATTTAGAGCAAG	601
Db	545	GGCATCTATCTAATTAAGAAATTAACGCTTTTAAACGCACAGAAATCAACCAATCAAT	604
Oy	602	AGATGGCGGATTAAGCCCTTTCTGCTTTTGGCCAGGCTTAGGACTTGTGTGCGCG	661
Db	605	TAAATAGTTAAAGTCTAGTAAATCAATGCAATGGCTGGAAGACAGAACAATGGCTTACCG	664
Oy	662	GGGATTTATACACGCCCTTCCCGGTTTGTGACAGCTGTTATTTGAAAGACATCTAAGC	721
Db	665	GTTGTTATATAGTACACAGATCGTTTATACGGCTACATATTTAAGACACCACTACGCT	724
Oy	722	CGGCGGCCGATGAACGAAAGGTATACACGCCGCTTTCAATTTTGGCAAAATATGACGA	781
Db	725	GTTTCCCATTAATGAAGATGAATAATTTATGATTAATGTTTTTAAAGTTCTAAGATCAGTCAGTA	784
Oy	782	TACCAAGGGCGCAGATGATACGGAAGAAAGCAAAATCATATACGATATATCTTCG	841
Db	785	TCCTTCAGGTGCAAGTTAT--CGATGCCAATTAATAATCATTTACACACATATCAATTAG	844
Oy	842	TGATGTGCAACGAAACTGCAAACTACTATTTCCACCATATGACAAATCGGCAAAATCCAAA	901
Db	842	TGATGGAAGTAAAGAAAGATTAATATATTAAGCTTATTTGCAATCAAAATTTCA	901
Oy	902	AAGTAATTTATTTCAATGA	920
Db	902	AAATTAATTTAAGTGAAGA	920

## RESULT 2

```

: Sequence 2421: Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 2421
: LENGTH: 996
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
: US-09-134-001C-2421

```

```

Query Match      16.0%;   Score 159;   DB 3;   Length 996;
Best Local Similarity 49.8%;   Pred. No. 1,4e-39;
Matches 458;   Conservative 0;   Mismatch 455;   Indels 6;   Gaps 2

QY      2  ATATGTGCACAGTCTTACATTCGGAACATCGCTGACCGCTAAACATGATATTGCAAGAACAA 61
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

D	14	AAATGTGACAGCCATTCTTTATATACAAAGCAAGTTACATTATTTAGTAGAACA	73
Q			
D	62	TGATTTTGCATTTCAGCTGGGGAACAAGTGATCTCTATCCGGCCGTTACAGCTGGA	121
Q			
D	74	TGCACTTTGCAATTTGAATTAATGTATCCCAACCATTTGTTCCAGCAATTTACATCC	133
Q			
D	122	ACAGTGAAGCTGACGGAAGGGCCCATCAAGACAAGTAACGCTTTATCGGTATGGGAGAA	181
Q			
D	134	AATTTGATCTAGATTCAAGCATGCGCTT--GAATATGTTTTTTGGAACAAAATTTAA	190
Q			
D	182	AACTTGAAAATATATTTATTTGCGACGCGCATTTATGAAAGCCGTTATCTGTGCGCGC	241
Q			
D	191	AAAGTAGACGTTATAGATTTGGTAGTGGATTAAGAAAAGAAAGSTTTAGATTTCCAAAC	256
Q			
D	242	TTTATTTTCCGGGCTATCGGGAATACGAAAAAAGATACGGAAGATACCGTTCAATTG	301
Q			
D	251	ATTACTTCACTGTGTAAGCCCTCATCAAGTACCATTAACGTTATGGTTATTTTAACTTAG	310
Q			
D	302	TCCGCAATGATTTTGTGACATGGGGGCTGTCAATCGTCAAGCTTTGGAAGACGTAAAG	361
Q			
D	311	CACCTGAAGAGTTTATTTGTTGGTTTAGGTTTAATTAAGATTTTACGAAATTAAC	370
Q			
D	362	AAAAGATTCGATCTTTAAGATTGTAGAGAAAAATATGATCTATTTGATACAGTTTAC	421
Q			
D	371	AAAGGTTAAGAAATCAATATTAATGAATGAAAAAATACGATTTGATATCGTTCTC	430
Q			
D	422	CGCTTCACTGATATTTGTCAAGCCGGAACGGGCGAAACCTGACGATGAACCAAGACAG	481
Q			
D	431	CTTTCATTTCAATGGTCACTGATGAACAGGACATACCGTACATAGAACCTCACAATG	490
Q			
D	482	ACGGCCTCAAAAGCTATGATATCAACTGGTGTATGACGGAACAGCCCCGATTATAT	541
Q			
D	491	GCTTATTAATAGTTAAAGATTAATTTATGTTCAATCACTTAACTTAATGAATGAACCTAAATTAAGATT	556
Q			
D	542	GGCATGTACCAATCTGACGACAAATTAACGAATCAGACCGAAGCAATTAGAGAGAAAG	601
Q			
D	551	GGCATCTATCTTACTTAAGAAATTAACGCTTTTAAACGCAACGAATACCAATCAAT	610
Q			
D	602	AGATGGGCGAATTAGCCCTTCTGCTTTTGAGCAAGGCTTAGAACTGTGGTGGCCG	661
Q			
D	611	TAAATGTAAAGTGTAGTAAGATCAATGGCGGTGGAACAGAAACAAATGGCTTACCGG	670
Q			
D	662	GGGATTAATACACCGGCTTCCGGTTTGTCAAGCTGTTATTTGAAAGAACATCTAGAGC	721
Q			
D	671	GTGGTTAATACGTCAACAATCGTTTATAACGGCTACATATTTAACAACCAACTACGCT	730
Q			
D	722	CGGGGCGCGATGAACGAAGGTGTAACAGCCGCTTTCAAAATTTGGCAAAATATGACGA	781
Q			
D	731	GTTCCCATTAATGAATGAATTAATTAATGAATTTGTTTAAAGTTCTTGAATCAGTCAATA	790
Q			
D	782	TACCAAAAGGCGCAGTGAATACGGAAGAAAGCAAGAAATCTATTATACGCAATATATCTCG	841
Q			
D	791	TCCCTCAAGGTGCAATTAT--CGATGCCAATAAATACATTACACATATCAATTAG	847
Q			
D	848	TGATGGAAGTAAAGAAAGAAAGTATTAATTAATTAAGCTTATTTGCAATCAAAATTTTCA	907
Q			
D	902	AAATTAATTTATTTTATGTA 920	
Q			
D	908	AAATTAATTTAATCTAAGA 926	

### RESULT 3

```

US-09-710-279-3345/c
; Sequence 3345, Application US/09710279
; Patent No. 6703492
;
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
;

```

PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 3345  
LENGTH: 2975  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-3345

Query Match 16.0%; Score 159; DB 3; Length 2975;  
Best Local Similarity 49.8%; Pred. No. 2.5e-39;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTGCAACAAGTCTTACATTTGGAACCTGCTAACCTTAACATGTATTGCAAGAACAA 61  
DB 1755 ATATGTGCTGCTGCTATTTTATATACAAACAGCTTACATTTATTTAGCTAGAACAA 1696  
QY 62 TGGATTTTGCATTTGAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 121  
DB 1695 TGGATTTTGCATTTGAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 1636  
QY 122 ACAGTAAGCTGACGGAAGGGCCCATCAGACAGTACCGGTTATCGGTATGGGGAGAA 181  
DB 1635 AATTGATCTAGATTCAGATCATGCGTCTT--GAATATGTTTGTGGAACAAATTAA 1579  
QY 182 AACTTGGAATATATATTGTCGCGAGGCTTAAAGAAAGGTTTATCTTGCGCGCGC 241  
DB 1578 AAGTAGAGCTTAAAGATTTGATGATGATTAACGAAAGGTTTATGATTTTCAAGC 1519  
QY 242 TTTATTTTCCGGGCTATGCGAGTACGAAACAAAGATACGGAAGATACCGTTACATTTG 301  
DB 1518 ATTACTTCACTGGTGAAGCTCATACGATCCATTAACGTTATGTTATTTTAACTTAG 1459  
QY 302 TCCCGCATAGTTTGAACATGGGTGCTGCTGCTGCTGCTGCTTGAAGAGCTAAAG 361  
DB 1458 CACCTGAGAGTTTATTTGTTGGTTTATTAATAAGATTTAGGATTTAAAC 1399  
QY 362 AAAAGTTTGCATTTTACGATTTGATGAGAAATTAAGATCTATTGATACGTTTAC 421  
DB 1398 AAAAGTTTGAATAATCAATTTATGATTAAGAAATTAAGATCTATTGATACGTTTAC 1339  
QY 422 CGCTTCACTGATATTTGTCAGACGCGGCGGAAACCTGACATGAAACGAAGCAG 481  
DB 1338 CTTTACATTTTCACTGATGTCATGATGAACGAGATACCGTACATTAACCTTCAATG 1279  
QY 482 ACGGCTCAAGTCTATGATTAATCAACTGCTGTCATGACGAACAGCCCGCATTTATAT 541  
DB 1278 GCTTATTAATAGTAAAGATTAATTTAGTTTATCTTAACTTAAGATTAATTTAGATT 1219  
QY 542 GGCATTAACCAATCTGACAGCAATATACAGAAATCAAGCCGAAGCAATTAGAGCAAG 601  
DB 1218 GGCATTAACCAATCTTAAAGAAATTAAGCTTTTAAAGCCACAGAAATCAACCAATCAAT 1159  
QY 602 AGATGGGCGGATTAAGCCCTTTGCTTTGGCCAAAGCTTAAAGATGTTGGTCCGCG 661  
DB 1158 TAAATAGTAAAGTCTAGTAAATCAATGAGCTGTAAGAGGAAACAAATGGCTTAAAC 1099  
QY 662 GGGATTTAAACAGCCCTTCCGCTTTGTGACAGCTGTTTATTTGAAGAAACATTTAGAC 721  
DB 1098 GTGGTTAAACGTCACAGATCGTTTATTAAGGCTTACATATTTAAAGCAACATACGCT 1039  
QY 722 CGGCGGCGCATGAAACGAAGGTATGACAGCCGCTTTCAATTTTGGCAATATAGCA 781  
DB 1038 GTTCCCAATATAGAAAGTAAATTTATGATTTTAAAGTTCTAGAGATCAAGTACAT 979  
QY 782 TACCAAAAGGCGCATGATTAACGGAAGAAAGCAAAATTTATTAAGCAATATTAATCTCG 841  
DB 978 TCCCTCAAGGTGCAATTAAT--CGATGCCAATTAATTAATTAACATTAATTAATTAAG 922

QY 842 TGATGTGCAACAAGTCTTACATTTGGAACCTGCTAACCTTAACATGTATTGCAAGAACAA 901  
DB 921 TGATGGAAGTAAAGAAAGATTTATTTATTAAGCTTACTTACATTAACAAATTTTCA 862  
QY 902 AAGTAAATTTATTTCTATGA 920  
DB 861 AATTAATAATTAAGTGAAGA 843

RESULT 4  
US-09-710-279-4184  
Sequence 4184, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P0348005  
CURRENT APPLICATION NUMBER: US/09/710,279  
PRIOR FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 4184  
LENGTH: 4114  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-4184

Query Match 16.0%; Score 159; DB 3; Length 4114;  
Best Local Similarity 49.8%; Pred. No. 2.9e-39;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTGCAACAAGTCTTACATTTGGAACCTGCTAACCTTAACATGTATTGCAAGAACAA 61  
DB 1626 ATATGTGCTGCTGCTATTTTATATACAAACAGTTACATTTATTTAGCTAGAACAA 1685  
QY 62 TGGATTTTGCATTTGAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 121  
DB 1686 TGGATTTTGCATTTGAGCTGGGACAGAGGTGATTCCTATCCGCGCGCTTACAGCTGGA 1745  
QY 122 ACAGTAAGCTGACGGAAGGGCCCATCAGACAGTACCGGTTATCGGTATGGGGAGAA 181  
DB 1746 AATTGATCTAGATTCAGATCATGCGTCTT--GAATATGTTTGTGGAACAAATTAA 1802  
QY 182 AACTTGGAATATATTTATTTGCGGACGCTTAATGAACGCGTTATCTTGCGCGCGC 241  
DB 1803 AAGTAGAGCTTATGATTTTGTGATGATTAAGAAAGGTTTATGCTATTTGCAAGC 1862  
QY 242 TTTATTTTCCGGGCTATGCGAGTACGAAACAAAGATACGGAAGATACGTTTACATTTG 301  
DB 1863 ATTACTTCACTGCTGTAAGGCTCATACAGTACCATTAACGTTATGTTTAACTTAG 1922  
QY 302 TCCCGCATAGTTTGAACATGGGTGCTGCTGCTGCTGCTGCTTGAAGAGCTAAAG 361  
DB 1922 CACCTGAGAGTTTATTTGTTGGTTTATTAAGTTTAAAGATTTAGCGAATTTAAAC 1982  
QY 362 AAAAGTTTGCATTTTAAAGATTTGATGAGAAATTAAGATCTATTTGATACATTTTAA 421  
DB 1983 AAAAGTTTGAATAATCAATTTATGAATGAABAAAATTAAGCATTTGAATATCGTTCCTC 2042  
QY 422 CGCTTCACTGATATTTGTCAGACGCGGCGGAAACCTGACATTAAGAACCAAGACAG 481  
DB 2043 CTTTACATTTTCAATGCTGATGTAAGAACAGATACCGTACCTATTAAGCTTCAATG 2102  
QY 482 ACGGCTCAAGTCTATGATTAATCAACTGCTGTCATGACGAACAGCCCGCATTTATAT 541  
DB 2103 GCTTATTAATAGTAAAGATTAATTTGTTTCACTTAAACAAATGAACCTTAATTTAGATT 2162  
QY 542 GGCATGTAACCAATCTGACGAAATTAACAGGAATCAAGCCGAAGCAATTTAGAGCAAG 601

```

Db      2163  GGCATCTATCTAATTAAGAAATTAACGCTTTTAAAGCCACAGAAATCAACCAATCAAT 2222
Qy      602  AGATGGCGGATTAAGCCCTTCTGCTTTGGCCAAAGCTTAGAACTGTGTGTCGCG 661
Db      2223  TAATAGTAAAGTCTAGTAAGATCAATGAGGCTGTGAAGCAGAAATGGCTTACCG 2282
Qy      662  GGGATTATACACCGCTTCCCGGTTTGTACAGCTGTTTATTTGAAAGAACATCTAGAG 721
Db      2283  GTGGTTATAGTCAACAGATCGTTTATAGCGCTACATATTTAAGACCACTAGCGT 2342
Qy      722  CGGCGCGCATGAACAAAGAGTGAACAGCGCTTTCAATTTGGCAATATGACGA 781
Db      2343  GTTCCCATTAATGAATGAATTAATTAATGATGTTTAAAGTCTAGAACCTGACGTA 2402
Qy      782  TACCAAGGCGGCGATTAACGAAAGACGAAATTCATTAAGCAATATCTCCG 841
Db      2403  TCCCTCAAGGTGCGATAT--CGATGCAATTAATAATACATTACACATATCAATTAG 2459
Qy      842  TGAATGCAACGAATGAAACTGAACTATTTCCACCACTATGACATCGGCAATCCAAA 901
Db      2460  TGAATGAAAGTAAAGAAAGATTTATATTAATTAAGCCTTACTTAGCATCAATTTTCA 2519
Qy      902  AAGTAAATTTATTTTCATGA 920
Db      2520  AATTAATTAATTAAGTAAAG 2538

```

RESULT 5  
US-08-956-171E-444/c  
Sequence 444, Application US/08956171E  
Patent No. 6593114

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5236

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 444:

SEQUENCE CHARACTERISTICS:

LENGTH: 1146 base pairs

TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 444:  
US-08-956-171E-444

Query Match 14.7%; Score 146; DB 3; Length 11466;  
Best Local Similarity 47.5%; Pred. No. 6,3e-35;  
Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

```

Qy      4  ATGTCACAAAGCTTACATGTAAGAACTGCTGACCGTAAACATGATATTAAGCAACAAATG 63
Db      3688  ATGTGACAGAGATTCACATACAAACTTTAAATACAACTACTCTTGGACGACAGATG 3629
Qy      64  GATTTGCAATTTGAGCTGGGAGACAGAGTATCTCTATCGCGCGCTTAAGCTGCAAC 123
Db      3628  GATTATGATATTCATTAGATAGTGTGCGCAGCAGTACGCGCTTAAGATTTCTTTGAAA 3569
Qy      124  AGTGAAGCTGACGGAAGGCGCCATCAGACACATAGCGCTTTATCGGTATGGGAGAAA 183
Db      3568  TCTTGCACCT--GCAACGACAGGCGCAACGCAATATGCTTTATGTCACAGAAACGAT 3512
Qy      184  CTTGAAATATATATTTATTTGCGACGCGCATTAATGAAAGCGTTTATCTTGTGCGCGCT 243
Db      3511  ATGAAAGCTTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3452
Qy      244  TATTTTCCGGGCTATGCGAGTAAGAAAAAGATACGGAAGATACCGTTACATTTGTC 303
Db      3451  TATTTCCGAGGTTATAGTTATCATATGATCAACACACAAAGCGGACGCAATATTTACG 3392
Qy      304  CCGCATGAGTTTGTGACATGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db      3391  CAATAAGTAATTTGTGACATGATGATTTGGATATACAAAGCATGTAAGATATGAAACAA 3332
Qy      364  AAGATTCATCTTTAAGATTTAGAGAAATTAATAGATCTATTTGATACATTTTACCG 423
Db      3331  CAAAGATCCCAATATCATGTTGATGCTGATATTTAATATGACATCGGTGAAGTTCCGCA 3272
Qy      424  CTTCATGATATTTGTCAGACCGGACCGGCGGAACTGACATGATACCAAGACGAC 483
Db      3271  TTGCATATCATGTTTCCGATGACATGACATGACATGACATGACATGATTTAAAGAGGT 3212
Qy      484  GGCCTCAAAGCTATGATATCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db      3211  GAACTGTTTAAAGATATCTTATGCTGCTTAAACAAATCAATCCAGACTTTAAATTTGG 3152
Qy      544  CATGTACCAATCTGACCAATATACAGATACAGACCGAAGCAATTTAGAGAGCAAG 603
Db      3151  CATATATGATATTTAAGACATATATATCAATATTTCTCTTATCCAGACACGAAATTTA 3092
Qy      604  ATGGCGGATTAAGCCCTTCTGCTTTTGGCCAAAGCTTAGAACTGTGCTGCGCGG 663
Db      3091  TTGAAAGGTGTAACGATTTGAACCTTTAGGCAATGACAGGTACATTTGGATTTCCAGGT 3032
Qy      664  GATTATACCGCTTCCCGGTTTGTGACAGCTGTTTATTTGAAGAAACATCTAGAGCG 723
Db      3031  GGATTTACTTCAACTGAGCGCTTGTGAGATGCAATTTAAGAAAGCAACATGCTCAA 2972
Qy      724  GCGGCGGATTAAGAAAGGTGTAACAGCGCTTTTCAAAATTTTGGCAATATGACGATA 783
Db      2971  AACCAATATTAAGAAAGATTTAATGATGATGATGATGATGATGATGATGATGATGAT 2912
Qy      784  CCAAGGCGCGCATGATACGGAAGAAAGCAAAATTCATTAACGCAATATATCTTCCGTG 843
Db      2911  CCGATGGAATTTGACGTCGCGCATGATGATGATGATGATGATGATGATGATGATGAT 2852
Qy      844  ATGTCACAAAGAACTGAAACTATTTTCCACATGACATTCGCGCAATCCAAATA 903
Db      2851  ATTAATTTAATCAAGAAAGTTATATTTAATTTAATTTAATTTAATTTAATTTAATTTA 2792
Qy      904  GTAATTTATTTCAATGAAAGCTTGTGCTGTTTGGAGCTTAAGTGTGTTTCCGCTAAAGCA 963
Db      2791  TTAAAGCTCAAGATGATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2732

```



QY 964 GAAGAGTATTCATGAGCTTAATTA 989  
DB 2731 CATATCACTATTAGAAAGTTGAATGA 2706

## RESULT 6

US-08-781-986A-444/C  
Sequence 444, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781.986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 444:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11466 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-444

Query Match 14.7%; Score 146; DB 3; Length 11466;  
Best Local Similarity 47.5%; Pred. No. 6.3e-35;

Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

QY 4 ATGTGCAAGTCTTACATGGAACCTGACCGTAACATGATTAAGCAACAATG 63  
DB 3688 ATGTGCAAGTCTTACATGGAACCTTAAATCAATCACTTCTTGACGACATG 3629  
QY 64 GATTTCGATTCAGCTGGGAGACAGAGTATCTTATCCGCGCTTACACTGGAC 123  
DB 3628 GATTATGATTAATCATTAAGTGTGCGCAGAGTACGCTGAATTAATCGTTGAAA 3569  
QY 124 AGTGAAGCTGACGAGAGGCGCCATCAGACACATGACCGCTTATCGGTATGGAGAAAA 183  
DB 3568 TCTTGACT--GGCAGCAGAGGCGCAACGCAATATGCGCTTATGCGCAGAGAAAGAT 3512  
QY 184 CTGGAATATATATATTTGCGCAGCATTAATGAAGCGTTATCTGTGCGGCGCTT 243  
DB 3511 ATGGAAGTATATATATATGATGATGATGATTAATGAACATGCGCTTGCATTTGAACAA 3452  
QY 244 TATTTTCCGGGCTATCGGAGTACGAAAAACGATACGGAGATACCGTTCACTTGTG 303  
DB 3451 TATTTCCGAGGTATATGTTTATGATCAACACACAAAGCGGACGATGAATTAACG 3392  
QY 304 CCGCATGAGTTTGTGACATGGGCTGTGCTGTGCTGCTGCTTGAAGACGTAAGAA 363

DB 3391 CAAATGAAATTTGACATGATTTTGGATATACAAACAGCATTTGAAGATATGAAACAA 3332  
QY 364 AAGATTCATCTTAAACATTTAGAGAAAAAATTAATGATCATATGATGATTAACCG 423  
DB 3331 CAAAGATCCCAATATACATGTTGTAGCTGTATATTTAATGATCGGTGAAGTTCCGCA 3272  
QY 424 CTTCATGATATATTTGACAGCGGAGCGGCGAAACCTGACGATTAAGAACAGACAC 483  
DB 3271 TTGCATTAATCATGTTTCCGATGCAATGACATACATGCAAGTTTCAATTAAGAGGT 3212  
QY 484 GGCTCAAACTATATATATCAACCTGTGTATGACGAAACGCCCCGACTTATATG 543  
DB 3211 GAAGTGTATTAAGATATCTTATGTTGTGCTTAAATCAATCAATCAATCAATTTGG 3152  
QY 544 CATGTAACCAATCTGACGATATACAGAAATGACGACCAAGCAATTAAGAGACAAAG 603  
DB 3151 CATTAATGATATTAATTAACATATATCAATATTTCTTATCAAGCAACGAAATTTA 3092  
QY 604 ATGGGCGGATTAAGCCCTTCTGTGCTTTGGCCAAAGCTTAAGAACTGTGTGCTGCCGG 663  
DB 3091 TTGGAAGGTGTAAGATTAAGCAATTAAGCAATTAAGCAATTAAGCAATTAAGCAAGT 3032  
QY 664 GATTATACAGCGCTTCCCGCTTTGTCAGAGCTGTTATTTGAAGAACATCTAGACCG 723  
DB 3031 GAATTAATCTTCACTGAGCGCTTTGGAAGTGGCAATTAAGAACCAATTTGCTCAA 2972  
QY 724 GCGGCGATGAAGAAAGGTGTAAGCGCGCTTTCAATTTTGGCAATTAAGACATA 783  
DB 2971 AACATGATTAAGAAATGATTTAATGATGATGATTTATTTAATGATGCGGTAAATTA 2912  
QY 784 CCAAGGCGGCGATGATTAAGCAAGAAAGCAAAATTCATTAACGAAATTAATCTCCGTG 843  
DB 2911 CCGATGGAATTTAGCTCGCGCATGATGCTGCAATCACTATGATGATGATGACCGTA 2852  
QY 844 ATGTGCAAGAACTGGAACCTATTTCCACCATATGACATGCGCAATCCAAAA 903  
DB 2851 ATAAATTTAATTAAGAAAGCTATATATTAATGATTAAGCAATGATTAAGTGA 2792  
QY 904 GTAATTTATTTATGATGACCTGACCTGTTGGAACCTTAAGTGTGTTCCGCTAAAGCA 963  
DB 2791 TTAAGCTCAGATATATTAATTAATGAAGAAATGATGATTTTAAGCTGAGAG 2732  
QY 964 GAAGAGATTAATCATGAGCTTAATTA 989  
DB 2731 CATATCACTATTAGAAAGTTGAATGA 2706

## RESULT 7

US-09-134-000C-838

Sequence 838, Application US/09134000C

Patent No. 6617156

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 032796-032

CURRENT APPLICATION NUMBER: US/09/134,000C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/055,778

PRIOR FILING DATE: 1997-08-15

NUMBER OF SEQ ID NOS: 6812

SOFTWARE: Patentin version 3.1

SEQ ID NO 838

LENGTH: 987

TYPE: DNA

ORGANISM: Enterococcus faecalis

US-09-134-000C-838

Query Match 9.6%; Score 95.4; DB 3; Length 987;

Best Local Similarity 45.8%; Pred. No. 1.6e-19;

Matches 407; Conservative 0; Mismatches 476; Indels 6; Gaps 2;  
QY 41 AACATGATTAAGCAAGAAACATGATTTTGCATTTCAAGCTGGGACAGAGGTGATTTCT 100



Db 605 GTTGAATGCTTAATAGCCGCGTATGAGGAGATAGGCTGCTGGAATTTATCTCAG 664  
QY 677 CTTCGCCGTTTGTGACAGCTGTTTATTTGAAGAATCATCTAGAGCCGCGCCGATGMAA 736  
Db 665 TATCTCGTTTGTAAAGCGACTTTTACGAGCTGAATCTGTATCTGAGATTCAGAGT 724  
QY 737 CGAAAGGTGACAGCCGCTTTTCAATTTGGCAATATATGACATACCAAGGCGCAG 796  
Db 725 CAGAAAGTATAGTCAATTTTTCATATCTTAGGTTAGAGAACACAAAGAGTTGT 784  
QY 797 TGATACGGAAGAGAGCAATATGATATACCAATATCTCCGTATGCAACGAA 856  
Db 785 GTATGTTGTTGAGGAAATATGAAATATACAAATTTATCTTGTTCATGTTGACA 844  
QY 857 CTGGAATCTATATTTCCACCATATGACATCCGCAATCCAAAGTAAATTTATTTTC 916  
Db 845 AAGGAATCTATATTTATGCAACATATGAGACATGCAATTTACTGCAATTTGATATGATA 904  
QY 917 ATGAAGCTTGAAGCTTTGGAGCCCTAAAGTGTTCGCTAAAGCAGAGAGATATTC 976  
Db 905 AAGAGGCTTATGATAGTATGATTAATGATATCAATATATGAAACAAACAAATTA 964  
QY 977 ATGAGCTTAATTA 989  
Db 965 AATATATTAATTA 977

## RESULT 9

US-08-956-171E-3720/c  
Sequence 3720, Application US/08956171E  
Patent No. 6593114

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439

## INFORMATION FOR SEQ ID NO: 3720:

SEQUENCE CHARACTERISTICS:  
LENGTH: 400 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 3720:  
US-08-956-171E-3720

Query Match 4.8%; Score 47.8; DB 3; Length 400;  
Best Local Similarity 48.1%; Pred. No. 0.0001;  
Matches 136; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 286 GATACCGTCAATTTGCGCGCATGAGTTTGTGACATGGGTGGTGTGAGTCTGACGTCT 345  
Db 333 GAGCCATGATATTTACGCAAAATGAAATTTGATGATGATTTTGGATATACACAAGC 274  
QY 346 TTGGAAGACGTAAAGAAAGATTCATCTTTAAGATTTAGAGAAATTAATGATCTA 405  
Db 273 ATGGAAGATATGAAACAAAGCATCCAAATATCATGTTTACTGTATATTTAAATGAC 214  
QY 406 TTGATATACATTTTACCGCTTCACTGATATTTGTGACACCGGACGCGCAAACTGACG 465  
Db 213 ATCGTGAAGTTCCGCGCATTTATCATGTTTCCGATGCACTGACATACAGTCGAA 154  
QY 466 ATGAACCAAGACAGACGCGCTCAAGTCTATGATATGCACTGCTGATGACGAC 525  
Db 153 GTTCAATTTAAAGAGGTGAAGTGTATTAAGATATATCTATTTGGTGTCTTAACAAT 94  
QY 526 AGCCCGACTTTATATGCGATGTACCAATCTGACGAATATA 568  
Db 93 CATCCAGACTTAATTTGGCATTTATGTAATTTAAGCAATATA 51

## RESULT 10

US-08-781-986A-3720/c  
Sequence 3720, Application US/08781986A  
Patent No. 6737248

## GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

## ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512

## INFORMATION FOR SEQ ID NO: 3720:

SEQUENCE CHARACTERISTICS:  
LENGTH: 400 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-3720

Query Match 4.8%; Score 47.8; DB 3; Length 400;

Best Local Similarity 48.1%; Pred. No. 0.0001;  
Matches 135; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 286 GATACCGTTCATGTCCTCCGATGAGTTGTGACATGGGTGCTGACGTCT 345  
Db 333 GACCGCATGATATATTACGCAAAATGAAATGTGACATGATTTGGGATATACAAAGC 274  
QY 346 TTGGAACACGTAAAGAAAAGATTCGATCTTTAAGATTGAGAAAATATGATCTA 405  
Db 273 ATGGAATATGAAACAAACGATCCCAATATACATGTTGATGATATTTAAATGAC 214  
QY 406 TTGGATACAGTTTACCGCTTCACTGATATTGTGACACCGAGCGCCGAAACCTGACG 465  
Db 213 ATCGTGAAGTTCGCCCATTCATTCATGTTCCGATGCACTGGACATACAGTGA 154  
QY 466 ATAGAACCAAGACGACGCGCTCAAAAGTCTATGATATCACTGGTGTATACGAC 525  
Db 153 GTTTCATTTAAAGAGGTGAGTGTATATAAAGATATCTTATGTTGTTTAACAAT 94  
QY 526 AGCCCGACTTATATGCGATGTAACCAATCTGACGACATATA 568  
Db 93 CATCGACCTAAATGGCATTAATGTAATTTAAGACATATA 51

## RESULT 11

US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEIFLINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESS: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: pTZpct-F15

US-08-232-463-14

Query Match 3.9%; Score 39; DB 2; Length 7218;  
Best Local Similarity 5.8%; Pred. No. 0.28;  
Matches 24; Conservative 206; Mismatches 181; Indels 0; Gaps 0;

QY 458 ACTGACATAGAACCAAGACGACGCTCAAAATCTATGATATCAACCTGGTGTCA 517  
Db 1436 ACRR 1377  
QY 518 TGACGACAGCCGACTTATATGATGATACCAATCTGACGACATATACAGATCA 577  
Db 1376 RRR 1317  
QY 578 GACCGAAGCAATTAAGAGACAGAGATGGCGATTAAGCTTTGCTTTGGCCAAG 637  
Db 1316 RRR 1257  
QY 638 GCTTAGAAGTGTGCTGCGGGGATTAACACCGCTTCCGGTTGTGACAGCTG 697  
Db 1256 RRR 1197  
QY 698 TTATATTGAAAGACATCTAGACCGCGCGATGAAAGAAAGGTGTACAGCCGCTT 757  
Db 1196 RRR 1137  
QY 758 TTCAAAATTTGGCAATATGACATACCAAGGCGCATGATACGAAAGAACGAAA 817  
Db 1136 RRR 1077  
QY 818 TTCATTATACGCAATATCTCCGTGATGCAAGCAAACTGGAATCTACT 868  
Db 1076 RRRRRRRRRATGCGAAGCTCCTCGACCTGCAAGCCGCAAGCTCGAATTAAT 1026

## RESULT 12

US-09-949-016-14834  
Sequence 14834, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14834

LENGTH: 107140

TYPE: DNA

ORGANISM: Human

US-09-949-016-14834

Query Match 3.8%; Score 37.4; DB 3; Length 107140;  
Best Local Similarity 49.7%; Pred. No. 3.7;  
Matches 95; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 682 CGTTTGTACAGAGCTTTATTTGAAAGACATCTAGAGCCGCGCGATGAACGAAA 741  
Db 3534 CGTTTGTGAGAGGTGATATTTAAAGGACAGAGCGAGCTTTAAAGAAA 3593  
QY 742 GGTGTACAGCGCTTTCAATTTTGGCAATATGACATACCAAGGCGGCTGTA 801  
Db 3594 AGGGAACAGAGAGATGTTTCTGTGAGGATGACCAAGAAAGACTGTGGGTCTT 3653  
QY 802 ACGGAAGAGAGCAATTCATTATACCAATATCTCCGTGATGTGCAAGAACTGA 861  
Db 3654 GCCATAGCACTTACCTTTGTAACGAAAACATCCCATTAATCATGATTTTGT 3713

OY 862 AACTACTATT 872  
Db 3714 AACTATT 3724

## RESULT 13

US-09-453-702B-62/C

Sequence 62, Application US/09453702B  
Patent No. 6365723

## GENERAL INFORMATION:

APPLICANT: Blatner, Frederick R.

Burland, Nicole T.

Punkett, Guy

Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles &amp; Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-Dec-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 61663

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-09-453-702B-62

Query Match 3.7%; Score 36.4; DB 3; Length 61663;

Best Local Similarity 49.5%; Pred. No. 5.7;

Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

OY 615 AGCCCTTTCGCTTTGGCCAAAGCTTGAAGCACTGTTGTCGCCGGGGATTATACACC 674  
Db 12938 AGCCCTTTCGCTTACCTCAAGCTTCGCAAGTTCAGGAGATGCCGACCTTTAATCGG 12879  
OY 675 GCGTCCCGGTTGTGAGAGCTGTTTATTGAAGAATCTAGAGCCGGGCGGATGA 734  
Db 12878 CTATCAAGATTTGATGAGATATCTTTTGTGACTTTCAGTACGTATATGTTGA 12819  
OY 735 AACGAAAGTGTAACAGCCGCTTTCAATTGGAATAATGAAGATACCAAGGCGCG 794  
Db 12818 ACTGAAAGTTTGGCAAGTGTGATGCTAGTACTAGTTCAGATGAAAGAGTGGCG 12759  
OY 795 AGTGATACG 804  
Db 12758 AAGAAATTCG 12749

## RESULT 14

US-10-114-170-62/C

Sequence 62, Application US/10114170  
Patent No. 685814

## GENERAL INFORMATION:

APPLICANT: Blatner, Frederick R.

Burland, Nicole T.

Punkett, Guy

Welch, Rod

TITLE OF INVENTION: No. 685814el Sequences of E. coli O157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles &amp; Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/114,170

FILING DATE: 01-Apr-2002

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/453,702

FILING DATE: 03-Dec-1999

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-Dec-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 61663

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 62:

US-10-114-170-62

Query Match 3.7%; Score 36.4; DB 3; Length 61663;

Best Local Similarity 49.5%; Pred. No. 5.7;

Matches 94; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

OY 615 AGCCCTTTCGCTTTGGCCAAAGCTTGAAGCACTGTTGTCGCCGGGGATTATACACC 674  
Db 12938 AGCCCTTTCGCTTACCTCAAGCTTCGCAAGTTCAGGAGATGCCGACCTTTAATCGG 12879  
OY 675 GCGTCCCGGTTGTGAGAGCTGTTTATTGAAGAATCTAGAGCCGGGCGGATGA 734  
Db 12878 CTATCAAGATTTGATGAGATATCTTTTGTGACTTTCAGTACGTATATGTTGA 12819  
OY 735 AACGAAAGTGTAACAGCCGCTTTCAATTGGAATAATGAAGATACCAAGGCGCG 794  
Db 12818 ACTGAAAGTTTGGCAAGTGTGATGCTAGTACTAGTTCAGATGAAAGAGTGGCG 12759  
OY 795 AGTGATACG 804  
Db 12758 AAGAAATTCG 12749



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocelebration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:52:29 ; Search time 992 Seconds  
(without alignments)  
8302.722 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996 1 catatgtgcacaaagctctac.....atgagcttaataagatcc 996

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications\_NA\_Main\*

1: /cgn2\_6/ptodaca/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodaca/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodaca/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodaca/1/pubpna/US10\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodaca/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodaca/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodaca/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodaca/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodaca/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodaca/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	996	100.0	996	9 US-10-812-387-1	Sequence 1, Appl1
2	439.8	44.2	990	3 US-09-974-300-1719	Sequence 1719, Ap
3	287.8	28.9	984	7 US-10-282-122A-9120	Sequence 9120, Ap
4	237	23.8	969	7 US-10-282-122A-9817	Sequence 9817, Ap
5	232.6	23.4	990	7 US-10-398-221-766	Sequence 766, App
6	232.6	23.4	990	7 US-10-398-221-2682	Sequence 2682, Ap
7	232.6	23.4	990	7 US-10-282-122A-24755	Sequence 24755, A
8	173.8	17.4	1616	6 US-10-429-802-17	Sequence 17, Appl1
9	173.8	17.4	1616	6 US-10-430-503-8	Sequence 8, Appl1
10	160.2	16.1	978	7 US-10-282-122A-34300	Sequence 34300, A
11	159	16.0	996	7 US-10-724-972A-1455	Sequence 1455, Ap
12	146	14.7	993	7 US-10-282-122A-1455	Sequence 7960, Ap
13	146	14.7	11466	2 US-08-781-986A-444	Sequence 444, App
14	146	14.7	11466	2 US-10-329-624-444	Sequence 444, App
15	142.8	14.3	1002	3 US-09-815-242-8383	Sequence 8383, Ap
16	142.4	14.3	981	3 US-09-815-242-4665	Sequence 4665, Ap
17	117.8	11.8	972	7 US-10-282-122A-35759	Sequence 35759, A
18	95.4	9.6	972	7 US-10-282-122A-20506	Sequence 20506, A
19	80.6	8.1	972	7 US-10-282-122A-21716	Sequence 21716, A
20	80.2	8.1	2256646	7 US-10-470-565-1	Sequence 1, Appl1
21	72.4	7.3	1050	7 US-10-282-122A-12199	Sequence 12199, A
22	63.6	6.4	978	7 US-10-398-221-911	Sequence 911, App
23	63.6	6.4	978	7 US-10-398-221-2826	Sequence 2826, Ap

24	61.8	6.2	579	7 US-10-282-122A-21539	Sequence 21539, A
25	55.2	5.5	1092	7 US-10-282-122A-40070	Sequence 40070, A
26	53.6	5.4	1092	7 US-10-282-122A-38691	Sequence 38691, A
27	52.2	5.2	9749	3 US-09-070-927A-154	Sequence 154, App
28	47.8	4.8	400	2 US-08-781-986A-3720	Sequence 3720, Ap
29	47.8	4.8	400	7 US-10-329-624-3720	Sequence 3720, Ap
30	40.8	4.1	37973	6 US-10-311-455-2169	Sequence 2169, Ap
31	37.2	3.7	1068	3 US-09-815-242-6788	Sequence 6788, Ap
32	37	3.7	10528	6 US-10-311-455-307	Sequence 307, App
33	36.4	3.7	61662	6 US-10-418-837-1	Sequence 1, Appl1
34	36.4	3.7	61663	5 US-10-114-170-62	Sequence 62, Appl1
35	36	3.6	826	3 US-09-070-927A-799	Sequence 799, App
36	35.8	3.6	583	3 US-09-925-065A-392869	Sequence 32869, A
37	35.8	3.6	3673778	6 US-10-312-841-2	Sequence 2, Appl1
38	35.4	3.6	245	3 US-09-080-107-1689	Sequence 1689, Ap
39	35.4	3.6	419	4 US-09-925-065A-654182	Sequence 654182, A
40	35.4	3.6	419	4 US-09-925-065A-654183	Sequence 654183, A
41	35	3.5	600	9 US-10-972-079-44650	Sequence 44650, A
42	35	3.5	961	7 US-10-648-593-102	Sequence 102, App
43	35	3.5	1053	9 US-10-450-763-28412	Sequence 28412, A
44	35	3.5	11036	5 US-10-239-676-118	Sequence 118, App
45	35	3.5	11036	6 US-10-240-453-138	Sequence 138, App

## ALIGNMENTS

RESULT 1  
US-10-812-387-1  
Sequence 1, Application US/10812387  
Publication No. US20050142652A1  
GENERAL INFORMATION:  
APPLICANT: SIVARAMAN, HEPHIZIAH  
APPLICANT: PUNDLE, ARCHANA VISHNU  
APPLICANT: SURBESH, CHERAVALAKATTU GOPALAN  
APPLICANT: DODSON, GEORGE GUY  
APPLICANT: BRANNIGAN, JAMES ANTHONY  
TITLE OF INVENTION: PROCESS FOR PRODUCTION OF LARGE AMOUNT OF PENICILLIN V  
FILE REFERENCE: 056859-0196  
CURRENT APPLICATION NUMBER: US/10/812,387  
CURRENT FILING DATE: 2004-03-30  
PRIOR APPLICATION NUMBER: PCT/IB03/06198  
PRIOR FILING DATE: 2003-12-24  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn Ver. 3.2  
SEQ ID NO 1  
LENGTH: 996  
TYPE: DNA  
ORGANISM: Bacillus subtilis  
US-10-812-387-1  
Query Match 100.0%; Pred. No. 1.6e-289; Length 996;  
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;  
Matches 996; Conservative 0;  
1 CATATGTGACAGCTCTTACATTGGAACCTGCTACCGTAAACATGATATTAGACAGAAC 60  
1 CATATGTGACAGCTCTTACATTGGAACCTGCTACCGTAAACATGATATTAGACAGAAC 60  
61 ATGATTTTGCATTTCAGCTGGGACAGAGCTATCTTATCCGGCCCTTACAGCTGG 120  
61 ATGATTTTGCATTTCAGCTGGGACAGAGCTATCTTATCCGGCCCTTACAGCTGG 120  
121 AATAGTGAAGCTGACGAGAGGGCCATCGACACACATGACGCTTATCCGATGAGGAG 180  
121 AATAGTGAAGCTGACGAGAGGGCCATCGACACACATGACGCTTATCCGATGAGGAG 180  
181 AATCTGGAATATATATTTTCCGACGACATTAATGAAGCGTTATCTTGTGCGCG 240  
181 AATCTGGAATATATATTTTCCGACGACATTAATGAAGCGTTATCTTGTGCGCG 240  
241 CTTTATTTTCCGGGCTATGCGAGATCGAAAAAAGCATACGAGAAATACCTTCACATT 300

DB 241 CTTTATTTTCCGGCTATGCGAGTACGAAAAAGATACGGGAAGATACGTTCACTT 300  
QY 301 GTCCCGCATGATTTGTGACATGGTGTCTGCACTGCTCTTGTGAAAGACGTAAA 360  
DB 301 GTCCCGCATGATTTGTGACATGGTGTCTGCACTGCTCTTGTGAAAGACGTAAA 360  
QY 361 GAAAGATTCATCTTTAACAATGTGTAGAGAAAAATAGATCTATGAGTACGTTT 420  
DB 361 GAAAGATTCATCTTTAACAATGTGTAGAGAAAAATAGATCTATGAGTACGTTT 420  
QY 421 CCGCTTCACTGATATTTGTACAGCCGAGCCGAACTGACGATGAAACCAAGACA 480  
DB 421 CCGCTTCACTGATATTTGTACAGCCGAGCCGAACTGACGATGAAACCAAGACA 480  
QY 481 GACGGCTCAAAAGTCTATGATATATCACTGTGTATGACGACAGCCCGACTTTA 540  
DB 481 GACGGCTCAAAAGTCTATGATATATCACTGTGTATGACGACAGCCCGACTTTA 540  
QY 541 TGGCATGTACCAATCTGACAGATATACAGGAATCAGACGAGCAATTAAGAGCAA 600  
DB 541 TGGCATGTACCAATCTGACAGATATACAGGAATCAGACGAGCAATTAAGAGCAA 600  
QY 601 GAGATGGCGGATTAAGCCCTTCTGCTTTTGGCCAGGCTTAAAGACTGTGCTGCG 660  
DB 601 GAGATGGCGGATTAAGCCCTTCTGCTTTTGGCCAGGCTTAAAGACTGTGCTGCG 660  
QY 661 GGGGATTTACACCCCTTCCGCTTTGTGACAGCTGTTTATTTGAAAGACATCTAAG 720  
DB 661 GGGGATTTACACCCCTTCCGCTTTGTGACAGCTGTTTATTTGAAAGACATCTAAG 720  
QY 721 CCGGCGGCGGATTAAGCAAGAGTGTAAACAGCCGCTTTCAATTTTGGCAATATGACG 780  
DB 721 CCGGCGGCGGATTAAGCAAGAGTGTAAACAGCCGCTTTCAATTTTGGCAATATGACG 780  
QY 781 ATACCAAGAGGCGCATGTATTAACGAGAGAGCAAAATTCATTATACGATATATCTCC 840  
DB 781 ATACCAAGAGGCGCATGTATTAACGAGAGAGCAAAATTCATTATACGATATATCTCC 840  
QY 841 GTGATGTGCAAGAACTGGAACCTATTTTCCACCATATGACAAATGGCAATCCAA 900  
DB 841 GTGATGTGCAAGAACTGGAACCTATTTTCCACCATATGACAAATGGCAATCCAA 900  
QY 901 AAGATTAATTTTATTCATGAGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTT 960  
DB 901 AAGATTAATTTTATTCATGAGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTT 960  
QY 961 GCAGAGAGAGTATTCATGAGCTTATTAAGATCC 996  
DB 961 GCAGAGAGAGTATTCATGAGCTTATTAAGATCC 996

RESULT 2  
US-09-974-300-1719  
Sequence 1719, Application US/09974300  
Patent No. US20020146721A1  
GENERAL INFORMATION:  
APPLICANT: Beirka, Randy M.  
TITLE OF INVENTION: Methods for Monitoring Multiple Gene  
FILE REFERENCE: 10085-500-US  
CURRENT APPLICATION NUMBER: US/09/974,300  
CURRENT FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: 09/680,598  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/279,526  
NUMBER OF SEQ ID NOS: 8481  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1719  
LENGTH: 990  
TYPE: DNA

ORGANISM: Bacillus licheniformis  
US-09-974-300-1719  
Query Match 44.2%; Score 439.8; DB 3; Length 990;  
Best Local Similarity 65.3%; Pred. No. 1.5e-121;  
Matches 645; Conservative 0; Mismatches 342; Indels 0; Gaps 0;  
QY 4 ATGTGCAAGATCTTATGATTAAGAACTGTGACCGGTAAACATGATTAAGCAAGAAATG 63  
DB 1 ATGTGCAAGATCTTATGATTAAGAACTGTGACCGGTAAACATGATTAAGCAAGAAATG 60  
QY 64 GATTTGCAATTTTCACTGAGGAGCAAGAGTATTTCTATCCGCCCTTATAGCTGAGAC 123  
DB 61 GATTTGCAATTTTCACTGAGGAGCAAGAGTATTTCTATCCGCCCTTATAGCTGAGAC 120  
QY 124 AGTGAAGCTGACGGAAGGAGCCATCAACACAGTACCGCTTATCGGATGAGGAGAA 183  
DB 121 AGTGAAGCTGACGGAAGGAGCCATCAACACAGTACCGCTTATCGGATGAGGAGAA 180  
QY 184 CTTGAAATATATATTTTGTGCGAGGCAATTAAGAAACGCTTATCTTGTGCGGCTT 243  
DB 181 CTTGAAATATATATTTTGTGCGAGGCAATTAAGAAACGCTTATCTTGTGCGGCTT 240  
QY 244 TATTTTCCGCGCTATGCGAGTACGAAACCAATACGGAAGTACCGTTCACTGTC 303  
DB 241 TATTTTCCGCGCTATGCGAGTACGAAACCAATACGGAAGTACCGTTCACTGTC 300  
QY 304 CCGCATGATTTGTGACATGAGGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCA 363  
DB 301 CCGCATGATTTGTGACATGAGGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCA 360  
QY 364 AAGATTCATCTTAAACATTTGAGAGAAATTAAGTCTATGATTAAGCTTATACG 423  
DB 361 GGGGCTGTTTCTTAAACATTTGAGAGAAATTAAGTCTATGATTAAGCTTATACG 420  
QY 424 CTTGATGATTTGTGACATGAGGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCA 483  
DB 421 CTTGATGATTTGTGACATGAGGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCA 480  
QY 484 GGGCTCAAGTCTATGATTAATCACTGTGCTATGAGCAAGCCCGCACTTATATG 543  
DB 481 GGGCTCAAGTCTATGATTAATCACTGTGCTATGAGCAAGCCCGCACTTATATG 540  
QY 544 CATGTAAACCAATCTGACAGCAATTAAGAAATCAGACCGAAACATTAAGAGCAAGAG 603  
DB 541 CATGTAAACCAATCTGACAGCAATTAAGAAATCAGACCGAAACATTAAGAGCAAGAG 600  
QY 604 ATGGGCGGATTAAGCCCTTCTGCTTTTGGCCAGGCTTGAAGACTGTGTGCGGCG 663  
DB 601 ATGGGCGGATTAAGCCCTTCTGCTTTTGGCCAGGCTTGAAGACTGTGTGCGGCG 660  
QY 664 GATTTATACCGGCTTCCGCTTGTGCAAGCTGTTATTAAGAAACATCTAGAGCGG 723  
DB 661 GATTTATACCGGCTTCCGCTTGTGCAAGCTGTTATTAAGAAACATCTAGAGCGG 720  
QY 724 GCGGCGGATTAAGCAAGAGGCTTCTGCTTTTCAATTTTGGCAATTAAGAGATA 783  
DB 721 GCGGCGGATTAAGCAAGAGGCTTCTGCTTTTCAATTTTGGCAATTAAGAGATA 780  
QY 784 CCAAGAGGCGGATTAAGCAAGAGGCTTCTGCTTTTCAATTTTGGCAATTAAGAGATA 843  
DB 781 CCAAGAGGCGGATTAAGCAAGAGGCTTCTGCTTTTCAATTTTGGCAATTAAGAGATA 840  
QY 844 ATGTGCAAGATCTGAAATCTATTTTCCACCTATGACATTCGCAATTCGCAATTC 903  
DB 841 ATGTGCAAGATCTGAAATCTATTTTCCACCTATGACATTCGCAATTCGCAATTC 900  
QY 904 GTTAAATTTATTCATGAGAACTTGAAGCTTGTGAGCTTAAAGTGTGTTCCGCTAAGCA 963  
DB 901 GTTAAATTTATTCATGAGAACTTGAAGCTTGTGAGCTTAAAGTGTGTTCCGCTAAGCA 960  
QY 964 GAAAGAGTATTCATGAGCTTATTA 990



Db 961 GAAGAGACGGTTCATTCTTCAATCA 987

## RESULT 3

US-10-282-122A-9120

Sequence 9120, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remainder Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 9120

LENGTH: 984

TYPE: DNA

ORGANISM: Bacillus anthracis

US-10-282-122A-9120

Query Match 28.9%; Score 287.8; DB 7; Length 984;

Best Local Similarity 56.0%; Pred. No. 1.2e-75;

Matches 544; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

Db 4 ATGTGCAAGCTTACATTTGAAATCTGACCGTAAATGATGATGAGAAACATG 63

1 ATGTGCAAGCTTACATTTGAAATCTGACCGTAAATGATGATGAGAAACATG 60

Qy 64 GATTTGCAATTCAGCTGGGAGACAGAGTATCTTATCCGCGCCGTTACAGCTGAG 123

Db 61 GACTTCACTTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120

Qy 124 AGTGAAGCTGACGAGAGGCGCATGACAGACGATCCGTTATCGTATGAGGAGAAA 183

Db 121 AATATTAACGGGTAAATCATTAATACGAAACATGCTACGATGCGATGATTAATCAT 180

Qy 184 CTGGAATATATATTTTGGCAGCGGATTAATGAAGCGTTATCTTGGCGGCGCT 243

Db 181 CAAGAGAGATCATTTATGCGGAGCGAGTAAATGACAGATGATGATGATGATGAT 240

Qy 244 TATTTTCGGGCTATGCGAGTACGAAAAACGATACGGAGATACCGTTACATTGTC 303

Db 241 TATTTTCAGAGATTCGCTACTTATATGATCAAGATACATGACACAGAAATTTGGCT 300

Qy 304 CCGATGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 363

Db 301 CCAATTTGATTTTGTAACTTGAAGCTGACCAATTTCAATTTCTGTCAAGATTTAAAGAA 360

Qy 364 AAGATTCGATCTTAAACGATTTGATGAGAAAAATATGATCTAATGATGATGATGAT 423

Db 361 TCTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

Qy 424 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483

Db 421 CTACATGATGATTTTGAAGGATTAATGAGGAGATGATGATGATGATGATGATGAT 480

Qy 484 GGCCTCAAGCTCTATGATTAATCAACCTGCTGATGATGATGATGATGATGATGAT 543

Db 481 GGAATTAATTTGATGATTAATCAACCTGATGATGATGATGATGATGATGATGAT 540

Qy 544 CATGTAACCAATCTGACGATATATACGATGATGATGATGATGATGATGATGATGAT 603

Db 541 CATTTACAAATTTTAAAGATATATATGATGATGATGATGATGATGATGATGATGAT 600

Qy 604 ATGGCGGATTAACCTTTCTGCTTTTGGCAAGGCTTAAAGAACTGTTGCTGCGGG 663

Db 601 TGAGATTAATTTACATTAATGCTTTTGGCAAGGCTGCGGGCTCAATGAGATTTCCAGGG 660

Qy 664 GATTAATACACCGCTTTCCGCTTTGATGATGATGATGATGATGATGATGATGATGAT 723

Db 661 GATTTACACCGCTTTCCGCTTTGATGATGATGATGATGATGATGATGATGATGAT 720

Qy 724 GCGGCGGATTAACCAAGGCTTAAACGCGCTTTTCAATTTTGGCAATTAATGATGAT 783

Db 721 ATGATAGCGAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

Qy 784 CCAAGGCGGCTGATTAACGAGAAAGCAAGAAATCATTAATGATTAATTTCTCGTG 843

Db 781 CTTAAAGGCTGATTAATTAACAGAAAGGCTGATTAATTAATTAATTAATTAATTAAT 840

Qy 844 ATGTGCAAGAACTGAAATCTTATTTCCACACATGATGATGATGATGATGATGAT 903

Db 841 ATGTGATGAAATCCGAAATTTTATTAATCAATTAATTAATTAATTAATTAAT 900

Qy 904 GTAATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963

Db 901 GTTCATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

Qy 964 GAAGAGATAT 974

Db 961 AATCAAAAAAT 971

RESULT 4

US-10-282-122A-9817

Sequence 9817, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See file Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 9817  
 LENGTH: 969  
 TYPE: DNA  
 ORGANISM: Bacillus anthracis  
 US-10-282-122A-9817

Query Match 23.8%; Score 237; DB 7; Length 969;  
 Best Local Similarity 55.5%; Pred. No. 2.5e-60;  
 Matches 538; Conservative 0; Mismatches 425; Indels 6; Gaps 4;

4 ATGTGACAAAGCTTACATGGAAAATGCTGACCGTAAACATGATTTAGCAAGAAATG 63  
 1 ATGTGACAAAGCTTACATGGAAAATGCTGACCGTAAACATGATTTAGCAAGAAATG 60  
 64 GATTTGATTTACAGCTGGGAGCAGAGGATTCCTCTTCCGGCCGTTACAGCTGAGAC 123  
 61 GACTTCACATTAAGATTAATCAAGAAATGATTAATCAATTCCTGACATTTAGCAAGAAATG 120  
 124 AGTGAAGCTGACGGAAGGCGCCATCAGACACAGTACGCGCTTATGCGATAGGGAAGAAA 183  
 121 AATATTAAGGCTGAATCATTAATTAAGCAACATGCTACGCTGGAAATGGATTTATCAT 180  
 184 CTGTGAAATATATTAATTTGCGCAGCGCATTAATGAAGCGGTTTATCTTTGCGCGCTT 243  
 181 CAAGGAAGATCATTAATGCGCAGCGAGTAATGAAGCGATGATGACATGTCACACATC 240  
 244 TATTTTCCGGGCTAATGCGGAGTAACAAAAACGATACGGAAGATACCGCTTACATGTC 303  
 241 TATTTTCCAGAGTTCGCTAATGCTAATGCTAAGCAATAGATGAACAAACGAATTTGGCT 300  
 304 CCGCATGATTTGTCAGATGGGCTGTCAGTCTGTCACTCTT--GGAAGACGTAAGA 361  
 301 CCAATTTGATTTGTAATCTTGAGTCTGACACATTCATTTCTGTCAAGGCCAGTGAAGA 360  
 362 AAAAGATTCGATCTTTAAGCATTTGAGAAAAAATTAGATCTATTGATAC-AGTTTAA 420  
 361 AATCTGATAGATGACATTAATCTTTTGGATATACATTAACCGCATTTAGACTTACGCCA 420  
 421 CCGCTTCACTGATATTTGTCAGACCGGAGCGCCGAAACCTGACGATTAACCAAGAGA 480  
 421 CCACATCAATTTGATTTTACCGGATTAATGCGGAGATTTGATGATGATCCGACAGT 480  
 481 GACGCGCTCAAGCTATGATATCAACTGCTGTGTCATGACGAACAGCCCGCATTTTAA 540  
 481 GAAAGATTTAAATTTGATGATTAACCACTAGAGATGATGACGAATGTCGGAATTTAA 540  
 541 TGGCATTTAACAATCTGACAGCAATATACAGAAATGACCGAAGCAATTTAGAGCAAA 600  
 541 TGGCATTTAACAATTTAGAGCAATATATATGAGCCTTAATTCGACCATTTGCGGCCAACA 600

601 GAGATGGCGGATTAAGCCCTTCTGCTTTTGSCCAAGCTTAGAATCTGTGCTGCGG 660  
 601 GAGTGAAGTAAATTAATCATTAATGCTTTTGCCCAAGCTCGGCTCAATGGGACTTCCA 660  
 661 GGGGATTTAATACCGCTTCCCGTTTGTGACAGCTGTTTATTTGAAAGAACATCTAGAG 720  
 661 GGGGATTTAATACCGCTTCCCGCTGAGGTTTGTGCGGACAGCAATATGCAAAACATTTCAA 720  
 721 CCGGCGGCGCATTAAGAAAGGATGTAACAGCCGCTTTTCAAAATTTGCAAAATATGACG 780  
 721 GGTATATATATGCGAAGAGGAGATGTCAGCCCTTTTCAATCTTATCAAAATTTGAG 780  
 781 ATACCAAGGCGGAGATATACGGAAGAGACGAATTCATTAATACCATATA-CTTC 839  
 781 GTTCTTAAGTGAAGTAAATTAACGAACAGGATGATTAATATACCATATATACCAAG 840  
 840 CGTATGTGCAACGAATCTGAAATCTATTTTCCACCACTATGACAAATCGG--CAATC 897  
 841 CGTATGTGATGGAATCCGGAACATATTTATCTTACTTACAAATGTAACCAATTA 900  
 898 CAATAAGTAAATTTATTCATGAAAGACCTTGACTTTGAGGCTTAAAGTTTTCGCT 957  
 901 TAGCTTGTCTATTTATTTCTGAAATTTAGATACCATGAAATTAAGCTTATCCCTTC 960  
 958 AAACGAGAA 966  
 961 CCACCGGAA 969

RESULT 5  
 US-10-398-221-766  
 Sequence 766, Application US/10398221  
 Publication No. US20040018514A1  
 GENERAL INFORMATION:  
 APPLICANT: KUNST, Frederick  
 APPLICANT: GLASER, Philippe  
 TITLE OF INVENTION: *Listeria innocua*, genome and applications  
 FILE REFERENCE: 344 702 - US  
 CURRENT APPLICATION NUMBER: US/10/398,221  
 CURRENT FILING DATE: 2003-03-27  
 PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
 PRIOR FILING DATE: 2001-10-04  
 PRIOR APPLICATION NUMBER: FR 00/12 697  
 PRIOR FILING DATE: 2000-10-04  
 NUMBER OF SEQ ID NOS: 4025  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO: 766  
 LENGTH: 990  
 TYPE: DNA  
 ORGANISM: *Listeria monocytogenes*-EGD  
 US-10-398-221-766

Query Match 23.4%; Score 232.6; DB 7; Length 990;  
 Best Local Similarity 53.0%; Pred. No. 5.5e-59;  
 Matches 496; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

4 ATGTGACAAAGCTTACATGGAAAATGCTGACCGTAAACATGATTTAGCAAGAAATG 63  
 1 ATGTGACAAAGCTTACATGGAAAATGCTGACCGTAAACATGATTTAGCAAGAAATG 60  
 64 GATTTGATTTACAGCTGGGAGCAGAGGATTCCTCTTCCGGCCGTTACAGCTGAGAC 123  
 61 GATTTGCTTTTATTTTGAAGGCGAACCAATATTAACCCCAAGAAATTAATGATGAGAAA 120  
 124 AGTGAAGCTGACGGAAGGCGCCATGACACAGTACGCGTTTATGCGATAGGGAAGAAA 183  
 121 TCTTCTACAGACGCGCAATTAATTAATGATATGATTTGAGCGGGAAGAGAG 180  
 184 CTGTGAAATATATTAATTTGCGACGCGATTAATGAAACCGGTTTATCTTTGCGCGCTT 243  
 181 TTAGTAAATATCAATTTTGCAGACGAGATTTGAATGAAGAAATTAAGCTGTGTCATCTCA 240

Oy 244 TATTTTCGGGGCTATGCGGAGTACGAAAAAACATACGGGAAGTACCGTTCATTTGC 303  
 Db 241 TATCTTCAGAGAAAGCATGTTATGCGCCACGACCACTGGAGGAAAAATCATATTTAGCT 3000  
 Oy 304 CCGCATGAGTTGTGACATGGGTGCTGCTCACTGTCACTCTTTGGAAAGCGTAAAGAA 3633  
 Db 301 CCACAAAGAAATTTTGGCTTTGGCTGTGTAGAACTTGGCCAAACGATTAAGATGTGGAAAGCA 3666  
 Oy 364 AAGATTCCGATCTTTACGATTTGAGAGAAAAATTTAGATCTATTTGATATACAGTTTACCG 4233  
 Db 361 AAGTTATCGGTTATTAACCTTGGAGATCAACCACTGATCCGTTACTTGAATTCACAAACCA 4200  
 Oy 424 CTTCACTGATATTTGACACCGGACGGGCCGAAACCTGACGATAGAACCAAGACGAC 4833  
 Db 421 TTACACTGGATTTTACAGACAAAGAGGGCGTTGTGTAGTATGATGAAACCAAGAAACA 4800  
 Oy 484 GGCCTCAAAGTCTATGATATATCACTCGTGTATGATGACAAACAGCCCGCATTTATATGG 5433  
 Db 481 TCGCTTCGTATTAAGAAAAAACCCCTGTAGAGATTATGCAAAATACGCCACGATCGAATGG 5400  
 Oy 544 CATGTAAACCAATCTGCAGCATATATACAGAAATGACCCGAAGCAATTAGAGCAAAAG 6033  
 Db 541 CATATGAGAGATTTAGCAACTACACAGGTTTACAGCAACACATTTAGGGCCGGTGAAG 6000  
 Oy 604 ATGGCGCGATTAGCCCTTTCTGCTTTTGGCCAAAGCTTAGAAACTGTTGCTTGGCCGGGG 6633  
 Db 601 TTCGAGAGATATATATGCTAAACATTTTCCGAAGGTACAGGAACAAGCAATTAACAGGT 6600  
 Oy 664 GATTTATACACCGGCTTCCCGGTTTGTACAGAGCTGTTTATTTGAAAGAACCTTAGAGCCG 7233  
 Db 661 GGTATATCTCCGCACAGACGTTTGTGTGTCGGCGCATCTTGAAAGAAAATATATTAATA 7200  
 Oy 724 GCGGCGGATGAAACGAAAGGTGTACAGCCGCTTTTCAATTTTGGCAATATGACGATA 7833  
 Db 721 GCGAAAAAGAGAAAGAGAGCTTACAAATGTGTGTATGTCTTGAACAGTGTCCGCAAT 7800  
 Oy 784 CCAAAAGGGCGAGTGAACGGAAGAAAGACGAATTCATTAATACGAATTAATCTTCGGTG 8433  
 Db 781 CCGAATGGACGGTATTTAAAGAAAGTGTGACCCCTGATTTTCAACAATATGTGGCACTCT 8400  
 Oy 844 ATGTGCACGAAACTTGAAACTATCTATTTTCCACACTATGACAACTGGCAATTCAGAAA 9033  
 Db 841 ATGTGTTCCGAAAGCAAAACGTACTATTTTCACTCTCATACGAAACAAACCAATATCAATAGT 9000  
 Oy 904 GTAAATTTATTCATGAAGACCTTGACTGTTTGGG 938  
 Db 901 GTTACTTGTGCAGAGAGATTATAGAAAATATGTAA 935

```

RESULT 6
US-10-398-221-2682
; Sequence 2682, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederick
; APPLICANT: GLASER, Philip
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 US
; CURRENT APPLICATION NUMBER: US/10/398, 221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2682
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Listeria monocytogenes EGDc
US-10-398-221-2682

```

Query Match 23.4%; Score 232.6; DB 7; Length 990;

Best Local Similarity	53.0%;	Pred. No. 5.5e-59;
Matches	496;	Mismatches 439; Indels 0; Gaps 0;

Oy	ATGGCAAGCTCTACATTGGAAATCGTCAACCGTAAACAGTATTTAGCAAGAACATG	63
Db	1 ATGTGCACAGCTTCGTATTAGAAACGTTAGATGAAACATTTATTATCAAGAACATG	60
Oy	64 GATTTTCATTCAGCTGGGGACAGAGTGATCTCTATCCGCGCGTTACAGCTGAAC	122
Db	61 GATTTTCCTTTATTATTGGAGGGCAACCAACATTTAGCCCAAGAAATTAATGATGAAA	12
Oy	124 AGTGAAGCTGACGGAAAGGGCCATCAGACACAGTACCGGTTTATCGGTATGGGGAGAAA	18
Db	121 TCTTCTACACAGCGGTCCCAATTATATTAAGATATGCAATTTGTTGAGCGGGAAAGAG	18
Oy	184 CTGGAAATATATTAATTTCGACGGCATTAATGAAAGCGGTTATCTTTGCGCGCTT	24
Db	181 TTAGATTAATATCAATTTTCGACAGCGATTGAATGAAGAAATTAAGCTGTGCATCTT	24
Oy	244 TATTTTCGGGCTATGCGAGTACGAAAAACGATACGGGAAGATACCGTTCACATGTC	30
Db	241 TATCTTCAGGAAGAACAGATTATATGCGCAGCACAGTGAAGAAATTAATTAAGCT	30
Oy	304 CCGCATGAGTTTGTGACATGGGTGCTGTCACTGTCACTCTTGGAAAGCTAAAGAA	36
Db	301 CCAAGAAATTTTGTGCTTGGCTGTATGAACTTCGCAAGATTAAGAAAGTGGAAACA	36
Oy	364 AAGATTCGATCTTTAAGATGAGTGAAGAAAAATTAGATCTATGAGTACAGTTTACG	42
Db	361 AAGTTATCGGTTATTAACCTTGGTGAATCAACAGTACCGTTACTTGGATACACACCA	42
Oy	424 CTTCACGTGATATGTCAAGCCGGAACGGGCGCAAACTGACGATGAAACCAAGACAGAC	48
Db	421 TTTCACGTGATTTTTCACACAAAGTGGGGTGTGTAGTATGATTGAACCAACGAACA	48
Oy	484 GGGCTCAAAAGTCTATGATTAATCAACCTGTGTGATGACGAACAGCCCGACTTATATG	54
Db	481 TCGCTTCGTATTAAGAAAAACCTGTGAGAGTTATGACAAATACGCCACGATCGAAATGG	54
Oy	544 CATGTAAACCAATCGCAGCATATACAGAAATCAGACCGAAGCAATTAGAGCAAGAG	60
Db	541 CATATGAGAAATTTATGCACTACAGAGTTTACAGCAACAAATTAAGCCCGGTGAAG	60
Oy	604 ATGGCGGATTAAGCCCTTCTGCTTTTGCGCAAGGCTTAGGAACGTGTGTCTGCGGG	66
Db	601 TTGCGAGAGTATATGCTTAACCATTTTTCGAGGTACAGGAACAAGAAATTAACAGT	66
Oy	664 GATTAATACACCGCTTCCCGGTTTGTCAAGCTGTTTATTGAAAGAACATCTAGACCG	72
Db	661 GGTATATCTCCGCCAGAAAGTTTGTCCGTGCGCATCTGAAAGAAATATATATAAA	72
Oy	724 GGCGCGAGTAACGAAAGGTGTACACGCGCTTTTCAAATTTTGGCAAAATATGACATA	78
Db	721 GCGAAAAAGAAAGAAAGAGCTATCACAATGTGTGTATGATCTTGAACAGTGTCCGAT	78
Oy	784 CCAAGAGGCGCAGTGTATAAGGAAGAAACGAAATTCATTATACGCAATATATCTCGTG	84
Db	781 CCGAATAGAGCGGTATTAAGAAAGTGTGACCTGATTTCAACACAAATATGTGGCATCT	84
Oy	844 ATGTGCACGAAATCTGAAATCTATATTTCCACCATATGCAATCGGCAATCCAAAA	90
Db	841 ATGTGTTCCGAAAGCAAAACGTAATTTTCACTCATTCGAAAAACAACCAATCAATAGT	90
Oy	904 GTAATTTTATTTATGAAAGACCTTGACGTTTGA	938
Db	901 GTTACTTGTCAAGAGTATATTAAGAAATATGAA	935

RESULT 7  
US-10-283-122A-24755  
; Sequence 24755, Application US/10282122A  
; Publication No. US80040029129A1  
; GENERAL INFORMATION:



```

Db      395 TTTACATGACGACAGATGTAAGTATATGTCCTCCAGTAATACGCAATTCGATG 454
Qy      127 GAAGCTGACGGA--GGGCCATCAGACACAGAACCGGTTATCGGTATGGGAGAA--181
Db      455 TTAGAAAAAGAAATGATGATTAACAATCATATGCTTTGTTGGAATGGAGAGACT 514
Qy      182 -AACTGGAAATATATATTTGCGGACGCAATTAAGAACGGTTATCTTGGCGGCG 240
Db      515 GACATTCATCACCAGATCTCTATGATGGGGTAAACGAAAGGAGATTAATGGGCGCATG 574
Qy      241 CTTATATTTCCGGGCTATGCGGAGTACGAAATAAGATACGGAGATACCGTTCACATT 300
Db      575 CTTTACTATGTAATTTCTCAAGTTTAGAAATGTGTACGTCGTGATGATGTTAT 694
Qy      301 GTCCCGATGATTTGTGACATGGGTGCTGTCACTGTCTGACGTTTGGAGACGTAAA 360
Db      635 AATCCGATGATGTAATTTCTCAAGTTTAGAAATGTGTACGTCGTGATGATGTTAT 694
Qy      361 GAAAAATTCGATCTTTAAGATTTGAGAAAAAATTAATCTATTTGATGATCACTTTTA 420
Db      695 GAAAAATTAATCTTTATACATTTATGATGAGCCATATATATCTGCTTTGCAACC 754
Qy      421 CCGCTTCACTGATATTTGTGACGCGGACGCGGCGGAAACCTGACGATAGAACGAAGCA 480
Db      755 CCACTTACATTAATTTACATTTAAGATGAGCCATATATATCTGCTTTGCAACC 814
Qy      481 GACGCTCAAGTCTATGATATCAACCTGTGTGATGACGAACAGCCGCACTTTATA 540
Db      815 ACAGGCAATTCATTCATTCGAAAAACATTTGGCGTCAATGAGATAGCCGTGATGAA 874
Qy      541 TGGATATTAACCATTCGACGACATATACGAAATCAGACGGAAGCAATTTAGACGAA 600
Db      875 TGGATATGACAAATTTAAGACTTATACATTTGTGTACACCAATTCGCCCAAGATATA 934
Qy      601 GAGATGGGCGGATTAAGCCCTTTCTGCTTTGGCCAGGCTTAGGAACGTTGTGCTGCG 660
Db      935 ATGATGGAGACTTGATTTGACACCGTTTGGGAGAGGGGCGAGGGGCTTAGGATTA 994
Qy      661 GGGGATTAACACCGCTTCCCGGTTTGTGACGCTGTTATTTGAAAGACATCTAGAG 720
Db      995 GGTGATTTTACGCGGTGAGCAGCTTTCTTGCGGTAGCATCTGAAAAAATATACGAA 1054
Qy      721 CCGGCGGCGGATGAAGAAAGGTGTACAGCCGCTTTCAATTTGGCAATATAGAG 780
Db      1055 AAAGCCAAATATGAACAGAAAGGCTTAACAATCTGTTCATATCTTCGTAAAT 1114
Qy      781 ATACCAAAAGGCGGATGATTAAGGAAAGAGAAATTCATTTATGCAATATATCTCC 840
Db      1115 ATCCCAAAAGGTGTTGTTTGAACAATGAGGGGAAACGGATTAATACATCTATACCTCA 1174
Qy      841 GTGATGTCAACGAACCTGAAACTATTTCCACGACTATGACAACTGGCAATCCAA 900
Db      1175 GCTATGTGACACAAATTAATACTATTTAACTGATATGACAAATGCGAATTTCA 1234
Qy      901 AAAGTAAATTTATTTATGAGACCTTGACTGTTTGAAGCTTAAAGCTTTTCCGCTAA 960
Db      1235 GCGGTTCTTAAAGGCTGAATTTAAATAGTCMAATTTAATTAATTAATGATGGAT 1294
Qy      961 GCAGAGAGATTTATGAGCTTAATTAAGA 993
Db      1295 CTTAAACAGATTTAAGCAATTAATCAAGTA 1327

```

## RESULT 9

```

US-10-430-503-8
; Sequence 8, Application US/10430503
; Publication No. US20040005684A1

```

```

; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: LAN, KENG-LI
; APPLICANT: OU-YANG, FU
; APPLICANT: LIU, JAM-CHING
; APPLICANT: LAN, KENG-HSIN

```

```

; TITLE OF INVENTION: TARGETING PROTEINS TO DELIVER THERAPEUTIC OR DIAGNOSTIC
; REAGENTS
; FILE REFERENCE: US/10430503
; CURRENT APPLICATION NUMBER: US/10430503
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/383,063
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1616
; TYPE: DNA
; ORGANISM: Bacillus sphaericus
US-10-430-503-8

```

```

Query Match      17.4%; Score 173.8; DB 6; Length 1616;
Best Local Similarity 49.8%; Pred. No. 4,1e-41;
Matches 495; Conservative 0; Mismatches 492; Indels 6; Gaps 2;

```

```

Qy      7 TGCACAACTTACATTTGAAACTGCTGACCGTAAACATGATTTAGCAAGAACATGAT 66
Db      335 TGCAGTAGCTTATCAATTCGTACACAGATGATTAAGTTATTCGCTCGCACATGAT 394
Qy      67 TTTGCATTTAGCTGGGAGACAGGTGATCTCTATCCGGCCGCTTACAGCTGAAACGT 126
Db      395 TTTACATGGAACCAATGATTAAGTATTTGTCCAGTAAATTAACGCAATTCGAT 454
Qy      127 GAAGCTGACGGA--GGGCCATCAGACACAGTACGCGTTATTCGATATGGGAGAA--181
Db      455 TTAGAAAAAGAAATGATGATTAACAATCATATGCTTTGTTGGAATGGAGAGACT 514
Qy      182 -AACTGGAAATATATATTTGCGGACGCAATTAAGAACGGTTATCTTGGCGGCG 240
Db      515 GACATTCATCACCAGATCTCTATGATGGGGTAAACGAAAGGAGATTAATGGGCGCATG 574
Qy      241 CTTATATTTCCGGGCTATGCGGAGTACGAAATAAGATACGGAGATACCGTTCACATT 300
Db      575 CTTTACTATGTAATTTCTCAAGTTTAGAAATGTGTACGTCGTGATGATGTTAT 694
Qy      301 GTCCCGATGATTTGTGACGCGGACGCGGCGGAAACCTGACGATAGAACGAAGCA 480
Db      755 CCACTTACATTAATTTACATTTAAGATGAGCCATATATATCTGCTTTGCAACC 814
Qy      421 CCGCTTCACTGATATTTGTGACGCGGACGCGGCGGAAACCTGACGATAGAACGAAGCA 480
Db      755 CCACTTACATTAATTTACATTTAAGATGAGCCATATATATCTGCTTTGCAACC 814
Qy      481 GACGCTCAAGTCTATGATATCAACCTGTGTGATGACGAACAGCCGCACTTTATA 540
Db      815 ACAGGCAATTCATTCATTCGAAAAACATTTGGCGTCAATGAGATAGCCGTGATGAA 874
Qy      541 TGGATATTAACCATTCGACGACATATACGAAATCAGACGGAAGCAATTTAGAGACAA 600
Db      875 TGGATATGACAAATTTAAGACTTATACATTTGTGTACACCAATTCGCAAGATATA 934
Qy      601 GAGATGGGCGGATTAAGCCCTTTCTGCTTTGGCAAGGCTTAGGAACGTTGTGCTGCG 660
Db      935 ATGATGGAGACTTGATTTGACACCGTTTGGGCAAGGGGCGGCGCTTAGGATTA 994
Qy      661 GGGGATTAACACCGCTTCCCGGTTTGTGACGCTGTTATTTGAAAGAACATCTAGAG 720
Db      995 GGTGATTTTACGCGGTGAGCAGCTTTCTTGCGGTAGCATCTGAAAAAATATACGAA 1054
Qy      721 CCGGCGGCGGATGAAGAAAGGTGTACAGCCGCTTTCAATTTGGCAATATAGAG 780
Db      1055 AAAGCCAAATATGAACAGAAAGGCTTAACAATCTGTTCATATCTTCGTAAAT 1114
Qy      781 ATACCAAAAGGCGGATGATTAAGGAAAGAGAAATTCATTTATGCAATATATCTCC 840

```



PRIOR APPLICATION NUMBER: 09/450,969  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: 09/134,001  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 7544  
SEQ ID NO 1455  
LENGTH: 996  
TYPE: DNA  
ORGANISM: S.epidermidis  
US-10-724-972A-1455

Query Match 16.0%; Score 159; DB 7; Length 996;  
Best Local Similarity 49.8%; Pred. No. 9.2e-37;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY 2 ATATGTCACAAAGCTTACATTTGAAACCTGTCACCGTAACATGTATTAAGCAAGAACAA 61  
DB 14 ATATGTCACAAAGCTTACATTTGAAACCTGTCACCGTAACATGTATTAAGCAAGAACAA 73  
QY 62 TGGATTTTGCATTTGACCTGGGACAGAGGTGATTCCTATCCCGCCGTTACAGCTGGA 121  
DB 74 TGGATTTTGCATTTGACCTGGGACAGAGGTGATTCCTATCCCGCCGTTACAGCTGGA 133  
QY 122 ACAGTGAAGTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181  
DB 134 AATTTGATCTAGATTCACAGATGCTGCTT--GAAATGCTTTGTTGAAACAAATTTAA 190  
QY 182 AACTTGGAAATATATTTTTCGCCAGCGCATTTAAGAAAGGTTTATCTTTGTCGGCGC 241  
DB 191 AAGTAGAGCGTTATAGATTTGATGATGATTAAGCAAGGTTTATGCTATTTCCAGAC 250  
QY 242 TTTATTTTCCGGGTATGCGGAGTACGAAAAAGATACGGAAGATACGCTTACATTTG 301  
DB 251 ATTACTTCACTGTGGAAGGCTCATACAGTACCCATTAACGTTATGTTTAACTTAG 310  
QY 302 TCCCGCATGATTTGTGACATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361  
DB 311 CACTGAGAGGTTTATGTTGGTTTATGTTTATTAAGATTAAGCAATTTAAAC 370  
QY 362 AAAAGATTCATCTTTAAGCATTTGTAGAGAAAAATTAAGATCTTATGATTAAGTTTAC 421  
DB 371 AAAAGATTAAGAAATCAATATTAAGATTAAGAAAAATTAAGATCTTATTAAGTTTAC 430  
QY 422 CGCTTCACTGATTTGTCAAGCGGAGCGGCGGAAACTGACGATTAAGCAAGCAAG 481  
DB 431 CTTTACATTTCACTGATTTGTCAAGCGGAGCGGCGGAAACTGACGATTAAGCAAGCA 490  
QY 482 AGCGGCTCAAGCTATGATTAATCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541  
DB 491 GCTTATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 550  
QY 542 GGCATTAACCAATCTGACCAATATACAGAAATCAAGCAAGCAAGCAATTAAGAGCAAG 601  
DB 551 GGCATTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 610  
QY 602 AGATGGCGGATTAAGCCCTTCTGCTTTTGGCCAGGCTTGAAGAACTGTTGCTGCGG 661  
DB 611 TTAATAGGTAAGGCTAGTAAAGTCAATGGGCTGTGAAGCAAGCAAGTAAAGTAAAGTAA 670  
QY 662 GGGATTAATACAGCGCTCCGGGTTGTGAGAGCTGTTTATTTGAAGAAAGTAAAGTAAAG 721  
DB 671 GTGCTTATACGTAACAGATGCTTTTATACGCGTACATATTTAAGCAAGCAAGTAAAGTAA 730  
QY 722 CGGCGGCGCAATGAAGAAAGTGTATACAGCGCTTTTCAATTTTGGCAATATATAGCA 781  
DB 731 GTTCCCATATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 790  
QY 782 TACCAAGGCGGAGTATTAAGCAAGCAAGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 841

DB 791 TCCCTCAAGGTGCAATTTAT--CGATGCCAATAAATACATTAACAAATATCATATAG 847  
QY 842 TGAATGCAAGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 901  
DB 848 TGAATGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 907  
QY 902 AAGTAAATTTATTTTCAATGA 920  
DB 908 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 926

RESULT 12  
US-10-282-122A-7960  
Sequence 7960, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITPA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7960  
LENGTH: 993  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
US-10-282-122A-7960

Query Match 14.7%; Score 146; DB 7; Length 993;  
Best Local Similarity 47.5%; Pred. No. 7.7e-33;  
Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

QY 4 ATGTGCAAGTCTTACATTTGAAACCTGACCGTAAACATGTATTAAGCAAGCAAGTAA 63  
DB 1 ATGTGCAAGTCTTACATTTGAAACCTGACCGTAAACATGTATTAAGCAAGCAAGTAAAG 60  
QY 64 GATTTGCAATTTGACCTGGGACAGAGGTGATTCCTATCCCGCGGCTTAAGCTGAAC 123  
DB 61 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120

124 AGTGAAGCTGACGGAAGGCCCATCAGACACAGTACCGCTTTATCGGTATGGGAGAAA 183  
121 TCTCGCACT--GGCAGACAGGCGCAACCGAATATGCTTTATGTTAGCAAGAAACGAT 177  
184 CTGGAAATATATATATTTGCGACGCGCATTAATGAAGCGGTTATCTTGTGCGCGCTT 243  
178 ATGGAAGCTTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237  
244 TATTTTCGGGCTATGCGGAGTACGAAAAAAGATACGGGAGATACCGTTACATTGTC 303  
238 TATTTTCGGGCTATGCGGAGTACGAAAAAAGATACGGGAGATACCGTTACATTGTC 297  
304 CCGCATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 363  
298 CAATATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 357  
364 AAGATTCGATCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 423  
358 CAGCATCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 417  
424 CTTCATGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483  
418 TTGCATTTATCATGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 477  
484 GGCTCAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543  
478 GAGTGTATTAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537  
544 CATGTAACCATCTGACGATATATACAGAAATGATGATGATGATGATGATGATGATGAT 603  
538 CATATATGATATTTAAGCAATATATATATATATATATATATATATATATATATATAT 597  
604 ATGGGCGGATTAAGCCCTTTCTGCTTTTGGCCAGGCTTATGAACTGTTGCTGCGGG 663  
598 TTGGAAGGTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657  
664 GATATATACAGCCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
658 GATATATACAGCCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717  
724 GCGGCGCATGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783  
718 AACATGATTAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
784 CCAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843  
778 CCGATGGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837  
844 ATGTCGAACGAAGCTGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 903  
838 ATAAATTTAATCTACAGAACTGATGATGATGATGATGATGATGATGATGATGATGAT 897  
904 GTAATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963  
898 TTAAGCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957  
964 GAGAGAGTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989  
958 CATATCACTATTAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 983

RESULT 13  
US-08-781-986A-444/c  
; Sequence 444, Application US/08781986A  
; Publication No. US20030054436A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5235  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville

STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781, 986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 444:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11466 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-444  
Query Match 14.7%; Score 146; DB 2; Length 11466;  
Best Local Similarity 47.5%; Pred. No. 2,9e-32;  
Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;  
4 ATGTCGAACGATCTTACATTTGAAAGCTGTCGACCGTTAAATGATGATGATGATGATGAT 63  
3688 ATGTCGAACGATCTTACATTTGAAAGCTGTCGACCGTTAAATGATGATGATGATGATGAT 3629  
64 GATTTTCATTTACGCTGGGGAAGAGGTGATTTCTATCCGCGCTTACGCTGGAAC 123  
3628 GATTTTCATTTACGCTGGGGAAGAGGTGATTTCTATCCGCGCTTACGCTGGAAC 3569  
124 AGTGAAGCTGACGGAAGGCCCATCAGACACAGTACCGCTTTATCGGTATGGGAGAAA 183  
3668 TCTTCGCACT--GGCAGACAGGCGCAACCGAATATGCTTTATGCGACAGAAACGAT 3512  
184 CTGGAAATATATATTTTGGCGACGCGATTAATGAAGCGGTTATCTTGTGCGCGCTT 243  
3511 ATGGAAGGTTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3452  
244 TATTTTCGGGCTATGCGGAGTACGAAAAAAGATACGGGAGATACCGTTACATTGTC 303  
3451 TATTTTCGGGCTATGCGGAGTACGAAAAAAGATACGGGAGATACCGTTACATTGTC 3392  
304 CCGCATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 363  
3391 CAATATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3332  
364 AAGATTCGATCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 423  
3331 CAGCATCCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3272  
424 CTTCATGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483  
3271 TTGCATTTATCATGTTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 3212  
484 GGCTCAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543  
3211 GAGTGTATTAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3152  
544 CATGTAACCATCTGACGATATATACAGAAATGATGATGATGATGATGATGATGATGAT 603  
3151 CATATATGATATTTAAGCAATATATATATATATATATATATATATATATATATATAT 3092



```

OY 604 ATGGCGGATTAAGCCCTTCTGCTTTGGCCAAAGCTTAGAAGCTGTGCTGCGGGG 663
DB 3091 TTGGAAAGGTGTACCATTAAGACCTTTAGGCAATGAAGCAATGATTTGATTTGCGAGGT 3032
OY 664 GATTATACCGCCCTTCCCGGTTTGTCCAGAGCTGTTTATTTAAAGAACCTTAAGCCG 723
DB 3031 GGATTTACTTCAACTGAGCGCTTTGTGAGATGCAATTTATGAAAGCAACCTTCTCAA 2972
OY 724 GCGGCGGATGAAGCAAGAGTGTAAACAGCCGCTTTTCAAAATTTGGCAAAATGACGATA 783
DB 2971 AACATATGAAGAAATGATTTAATGAATGATTTTATTTATGATCGGTAATATTA 2912
OY 784 CCAAGGCGCGAGTATTAACGGAAGACGAATTCATTATAGCAATATACCTTCGCTG 843
DB 2911 CCGATTGAATTTGTACGTCGCGCATATGCTGACATCACTATACGATGATCAGACCGTA 2852
OY 844 ATGTGCAACGAACCTGAAACCTATTTCCACCATATGACAAATCGGCAATCCAAAA 903
DB 2851 ATAAATTTAACTACAAAGACCTTATATATTAATTAATGCGCAAGCAATGATTAAGTACA 2792
OY 904 GTAAATTTATTTCAAGAACCTTGACTGTGAGCCCTAAAGTGTTCGCTAAAGCA 963
DB 2791 TTAAAGCTCAGATGATTTATTAATTAATGAAGAAATGATGAGATTTTAAAGCTTGAAG 2732
OY 964 GAAGAGATTTATCATGAGCTTAATTA 989
DB 2731 CATATCACTATTAAGAAAGTTGATGA 2706

```

RESULT 14  
 US-10-329-624-444/c  
 ; Sequence 444, Application US/10329624  
 ; Publication No. US20040043037A1  
 ; GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 Gil H. Choi  
 Patrick S. Dillon  
 Craig A. Rosen  
 Steven C. Barash  
 Michael R. Fannon  
 TITLE OF INVENTION: *Staphylococcus aureus* Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5256  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/329,624  
 FILING DATE: 27-Dec-2002  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/956,171  
 FILING DATE: October 20, 1997  
 APPLICATION NUMBER: 60/009,861  
 FILING DATE: January 5, 1996  
 APPLICATION NUMBER: 08/781,986  
 FILING DATE: January 3, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mark J. Hyman  
 REGISTRATION NUMBER: 46,789  
 REFERENCE/DOCKET NUMBER: PB248P1D1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (240) 314-1224  
 TELEFAX: (301) 309-8439  
 INFORMATION FOR SEQ ID NO: 444:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 444:
US-10-329-624-444

Query Match      14.7%; Score 146; DB 7; Length 11466;
Best Local Similarity 47.5%; Pred. No. 2.9e-32;
Matches 468; Conservative 0; Mismatches 515; Indels 3; Gaps 1;

OY 4 ATGTGCAACAGCTTAACTTGAAGAAATGCTGACCGGTAAACATGATTAAGCAAGAAATG 63
DB 3688 ATGTGCAACAGATTAACAATTAACAACTTTAAATATCAAGATATCTTGAACGACGATG 3629
OY 64 GATTTCGATTTTACGCTGGGACAGAGTGATTTCTATCCGCCCGCTTAAGCTGGAAC 123
DB 3528 GATTATGATTAATCAATTAAGATGTTGCCAGACAGTACGCTTAAGAAATTAATCGTTGGAAA 3569
OY 124 AGTGAAGCTGACGGAAGGCGCCATCAGACACAGTACGCGTTATCGGTAATGCGGAGAAA 183
DB 3568 TCTTGCACT--GGACAGACAGCGCAACGAATATGCTTTATTTGGCAGACAGACAT 3512
OY 184 CTTGAAATATATTAATTTTCCGACGCGCATTAATGAAGCGGTTTATCTTGTGCGCGCTT 243
DB 3511 ATGGAAGTTTATTAATTAATGATGATGTTATGAATGAATGACGTCGTCATTTCAACAAA 3452
OY 244 TATTTCCGCGCTATGCGGAGTACGAAAAAAGATAGGGAAGATACCGTTCACTTGTCTC 303
DB 3451 TATTTCCGAGGTATTAATTCATATGATGATCAACACAAAGCGACGATAAATTAATGCG 3392
OY 304 CCGCATGATTTTGAATGATGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
DB 3391 CAATATGAATTTGATGATGATTTTGGATATCAACAGCATTTGAAGATATGAACAA 3332
OY 364 AAGATTCGATCTTTAAGATTTGAGAAAAAATTAATTAATTAATTAATTAATTAATTAAT 423
DB 3331 CAAGCATCCCAATATACATGTTGTGATGATTAATTAATTAATTAATTAATTAATTAATTA 3272
OY 424 CTTACATGATTTGTGACAGCCGACCGGCGGAAACCTGACATGAAACCAAGACAGAC 483
DB 3271 TTGCATTAATCATGTTTCCGATGACATGACATGACATGACATGACATGACATGACATG 3212
OY 484 GGCTTCAAGCTATGATTAATCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB 3211 GAAGTGTATTAAGAAATATCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3152
OY 544 CATGTAACCAATCTGACAGCAATATACAGATCAAGCCGAAGCAATTAAGAGCAAAAG 603
DB 3151 CATATATGTAATTTAAGACATATATATCAATATTTCTCTTATCCAGAACAGCAAAATTTA 3092
OY 604 ATGGCGGATTAAGCCCTTCTGCTTTTGGCCAAAGCTTAGAAGCTGTGCTGCGGGG 663
DB 3091 TTGGAAAGGTGTACCATTAAGACCTTTAGGCAATGAAGCAATGATTTGATTTGCGAGGT 3032
OY 664 GATTATACCGCCCTTCCCGGTTTGTCCAGAGCTGTTTATTTAAAGAACCTTAAGCCG 723
DB 3031 GGATTTACTTCAACTGAGCGCTTTGTGAGATGCAATTTATGAAAGCAACCTTCTCAA 2972
OY 724 GCGGCGGATGAAGCAAGAGTGTAAACAGCCGCTTTTCAAAATTTGGCAAAATGACGATA 783
DB 2971 AACATATGAAGAAATGATTTAATGAATGATTTTATTTATGATCGGTAATATTA 2912
OY 784 CCAAGGCGCGAGTATTAACGGAAGACGAATTCATTATAGCAATATACCTTCGCTG 843
DB 2911 CCGATTGAATTTGTACGTCGCGCATATGCTGACATCACTATACGATGATCAGACCGTA 2852
OY 844 ATGTGCAACGAACCTGAAACCTATTTCCACCATATGACAAATCGGCAATCCAAAA 903
DB 2851 ATAAATTTAACTACAAAGACCTTATATATTAATTAATGATGATTTTAAAGCTTGAAG 2792
OY 904 GTAAATTTATTTCAAGAACCTTGACTGTGAGCCCTAAAGTGTTCGCTAAAGCA 963

```

Db 2791 TTAAGCTCAGAGATTTAATTAATAGAAAGATGACGATTTTAAGCTGAGAG 2732  
 Qy 964 GAAGAGATATTCATGACCTTAATTA 989  
 Db 2731 CATATCACTATTAGAAAGTTGAATGA 2706

RESULT 15

US-09-815-242-8383  
 / Sequence 8383, Application US/09815242  
 / Patent No. US20020061569A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Haselbeck, Robert  
 / APPLICANT: Ohlsen, Karl L.  
 / APPLICANT: Zykkind, Judith W.  
 / APPLICANT: Wall, Daniel  
 / APPLICANT: Trawick, John D.  
 / APPLICANT: Carr, Grant J.  
 / APPLICANT: Yamamoto, Robert T.  
 / APPLICANT: Xu, H. Howard  
 / TITLE OR INVENTION: Identification of Essential Genes in  
 / TITLE OF INVENTION: Prokaryotes  
 / FILE REFERENCE: ELITRA.011A  
 / CURRENT APPLICATION NUMBER: US/09/815,242  
 / PRIOR FILING DATE: 2001-03-21  
 / PRIOR APPLICATION NUMBER: 60/191,078  
 / PRIOR FILING DATE: 2000-03-21  
 / PRIOR APPLICATION NUMBER: 60/206,848  
 / PRIOR FILING DATE: 2000-05-23  
 / PRIOR APPLICATION NUMBER: 60/207,727  
 / PRIOR FILING DATE: 2000-05-26  
 / PRIOR APPLICATION NUMBER: 60/242,578  
 / PRIOR FILING DATE: 2000-10-23  
 / PRIOR APPLICATION NUMBER: 60/253,625  
 / PRIOR FILING DATE: 2000-11-27  
 / PRIOR APPLICATION NUMBER: 60/257,931  
 / PRIOR FILING DATE: 2000-12-22  
 / PRIOR APPLICATION NUMBER: 60/269,308  
 / PRIOR FILING DATE: 2001-02-16  
 / NUMBER OF SEQ ID NOS: 14110  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO 8383  
 / LENGTH: 1002  
 / TYPE: DNA  
 / ORGANISM: Staphylococcus aureus  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (1)...(1002)  
 / US-09-815-242-8383

Query Match 14.3%, Score 142.8; DB 3; Length 1002;  
 Best Local Similarity 47.3%; Pred. No. 7.2e-32;  
 Matches 466; Conservative 0; Mismatches 517; Indels 3; Gaps 1;

Qy 4 ATGTGCAAGCTTCTTATGTAAGTCTGACCGTAAACATGATTAATGCAAGAACATG 63  
 Db 10 ATGTGCAAGCTTCTTATGTAAGTCTGACCGTAAACATGATTAATGCAAGAACATG 69  
 Qy 64 GATTTGCACTTTCAGCTGGGAGACAGAGTGATTTCTATCCGCGCCGTTACAGCTGGAAC 123  
 Db 70 GATTAATGATTAATTCATTAATGATGTTGCGACAGTACCGCTTAATTAATGATGATG 129  
 Qy 124 AGTGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 183  
 Db 130 TCTGCGACT--GGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 186  
 Qy 184 CTGGAATATATATTTATTTGCGAGCGGATTAATGAAGCGGTTATCTTGTGCGGCGCTT 243  
 Db 187 ATGGAAGCTTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 246  
 Qy 244 TATTTCCGGGCTATGCGGAGTACGAAAAAGATGCGGAGATACCGTTACATTTGTC 303

Db 247 TATTCGAGGTATTAAGTCTCATATGCTCAACACACAAAGCGGATGATTAATTCG 306  
 Qy 304 CCGATGATTTTGTGACATGAGGTGCTGCACTGTGCACTTTTGAAGAGTAAAGAA 363  
 Db 307 CAATATGATTTGTGACATGATTTTGGATTAATCAACAGCATTTGAAGATGAAACA 366  
 Qy 364 AAGATTCGATCTTAAAGATTTGAGAAAAAATTAATGATCTTATGATGATGATTTACG 423  
 Db 367 CAAGCATCCCAATATCATGTTGTGATGATGATGATGATGATGATGATGATGATGATG 426  
 Qy 424 CTTCATGATGATTTGTGACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 483  
 Db 427 TTGATTAATCAATGTTTCCGATGCAATGACATTCAGTTCAGTTTAAAGAGGT 486  
 Qy 484 GGCCTCAAGCTTATGATTAATCAACCTGATGATGATGATGATGATGATGATGATGATG 543  
 Db 487 GAATGATTAATTAAGATTAATCAATGATGATGATGATGATGATGATGATGATGATGAT 546  
 Qy 544 CATGTAACCAATCTGACGATTAATCAAGATTAATCAAGATTAATCAAGATTAATCAAG 603  
 Db 547 CATTAATGATTAATTAAGATTAATCAATTAATCAATTAATTAATTAATTAATTAATTA 606  
 Qy 604 ATGGCGGATTAAGCTTTCGCTTTGGCCAGGCTTAGGAATGATGATGATGATGATGATG 663  
 Db 607 TTGAAGGTGAAGATTAATCAATTAATCAATTAATCAATTAATCAATTAATCAATTA 666  
 Qy 664 GATTAATCAAGCTTTCGCTTTGGCCAGGCTTAGGAATGATGATGATGATGATGATGAT 723  
 Db 667 GGATTAATCAAGCTTTCGCTTTGGCCAGGCTTAGGAATGATGATGATGATGATGATGAT 726  
 Qy 724 GCGGCGGATTAAGCTTTCGCTTTGGCCAGGCTTAGGAATGATGATGATGATGATGAT 783  
 Db 727 AACATGATTAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 786  
 Qy 784 CCAAGGCGGATTAAGCTTTCGCTTTGGCCAGGCTTAGGAATGATGATGATGATGATGAT 843  
 Db 787 CCGATTAATGATTAAGCTTTCGCTTTGGCCAGGCTTAGGAATGATGATGATGATGATGAT 846  
 Qy 844 ATGTGCAAGCTTCTTATGTAAGTCTGACCGTAAACATGATTAATGCAAGAACATG 903  
 Db 847 ATTAATTTAATCAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 906  
 Qy 904 GTAAATTTATTTCAATGAAGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 963  
 Db 907 TTAAGCTCAGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 966  
 Qy 964 GAAGAGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 989  
 Db 967 CATATCACTATTAAAGTTGAATGA 992

Search completed: March 18, 2006, 13:28:30  
 Job time : 998 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2006, 11:54:08 ; Search time 452 Seconds  
(without alignments)  
5138.006 Million cell updates/sec

Title: US-10-812-387-1

Perfect score: 996  
Sequence: 1 catatgtgcacaaagctctac.....atgagcttaataaagatcc 996

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 8023312 seqs, 1165852854 residues

Total number of hits satisfying chosen parameters: 16046624

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

Published Applications\_NA\_New.\*  
1: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/PCR\_NEW\_PUB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.1.\*  
7: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.1.\*  
9: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.1.\*  
10: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.2.\*  
11: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.3.\*  
12: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.4.\*  
13: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	16.0	980	US-10-793-626-413	Sequence 413, App
2	159	16.0	2975	US-10-793-626-3345	Sequence 3345, Ap
3	159	16.0	4114	US-10-793-626-4184	Sequence 4184, Ap
4	93	9.3	975	US-11-074-176-367	Sequence 367, App
5	83.8	8.4	975	US-11-074-176-369	Sequence 369, App
6	38.8	3.9	720	US-10-750-623-59863	Sequence 59863, A
7	38.8	3.9	720	US-10-750-623-59863	Sequence 59863, A
8	36.2	3.6	1272	US-11-096-568A-29311	Sequence 29311, A
9	35.8	3.6	583	US-09-925-065A-392869	Sequence 392869, A
10	35.4	3.6	419	US-09-925-065A-654182	Sequence 654182, A
11	35.4	3.6	419	US-09-925-065A-654182	Sequence 654182, A
12	35.4	3.6	5901	US-10-932-182A-2969	Sequence 2969, Ap
13	35.4	3.6	5901	US-10-932-182A-2969	Sequence 2969, Ap
14	35	3.5	961	US-11-072-175-102	Sequence 102, App
15	34.6	3.5	583	US-09-925-065A-392868	Sequence 392868, A
16	34.4	3.5	4429	US-10-753-195-9	Sequence 9, Appl1
17	34.2	3.4	583	US-09-925-065A-392870	Sequence 392870, A
18	34.2	3.4	1835	US-09-925-065A-45229	Sequence 45229, A
19	34.2	3.4	1835	US-09-925-065A-45229	Sequence 45229, A
20	34.2	3.4	1835	US-09-925-065A-45230	Sequence 45230, A

21	34	3.4	601	US-09-925-065A-348226	Sequence 348226, A
22	34	3.4	193084	US-11-121-086-82	Sequence 82, Appl1
23	33.8	3.4	670	US-09-925-065A-556802	Sequence 556802, A
24	33.6	3.4	556	US-09-925-065A-155341	Sequence 155341, A
25	33.6	3.4	556	US-09-925-065A-777072	Sequence 777072, A
26	33.4	3.4	572	US-09-925-065A-29870	Sequence 29870, A
27	33.4	3.4	580	US-09-925-065A-321496	Sequence 321496, A
28	33.4	3.4	653	US-09-925-065A-96912	Sequence 96912, A
29	33.4	3.4	659	US-09-925-065A-510272	Sequence 510272, A
30	33.4	3.4	958	US-09-925-065A-713894	Sequence 713894, A
31	33.4	3.4	958	US-09-925-065A-713895	Sequence 713895, A
32	33.4	3.4	1187	US-10-750-185-62875	Sequence 62875, A
33	33.4	3.4	1187	US-10-750-185-62875	Sequence 62875, A
34	33.4	3.4	1217	US-09-925-065A-668875	Sequence 668875, A
35	33.4	3.4	2830	US-09-925-065A-669245	Sequence 669245, A
36	33.4	3.4	2830	US-09-925-065A-669246	Sequence 669246, A
37	33.2	3.3	629	US-09-925-065A-346097	Sequence 346097, A
38	33.2	3.3	645	US-09-925-065A-860511	Sequence 860511, A
39	33.2	3.3	961	US-09-925-065A-690516	Sequence 690516, A
40	33.2	3.3	2037	US-10-932-182A-1230	Sequence 1230, Ap
41	33.2	3.3	2037	US-10-932-182A-1230	Sequence 1230, Ap
42	33.2	3.3	156297	US-11-121-086-65	Sequence 65, Appl1
43	33	3.3	393	US-09-925-065A-148412	Sequence 148412, A
44	33	3.3	499	US-09-925-065A-845854	Sequence 845854, A
45	33	3.3	566	US-09-925-065A-828854	Sequence 828854, A

# ALIGNMENTS

RESULT 1  
US-10-793-626-413  
Sequence 413, Application US/10793626  
Publication No. US20050255478A1  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
FILE REFERENCE: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
CURRENT APPLICATION NUMBER: US/10/793, 626  
CURRENT FILING DATE: 2004-03-04  
PRIOR APPLICATION NUMBER: 60/164, 258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 413  
LENGTH: 990  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-413

Query Match 16.0%; Score 159; DB 8; Length 990;  
Best Local Similarity 49.8%; Pred. No. 2e-35;  
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

QY	2	ATATGTCACAAAGTCTTACATGGAAGCTGACCGCTTAACATGATAGCAAGAACAA	61
DB	8	ATATGTCACAAAGTCTTACATGGAAGCTGACCGCTTAACATGATAGCAAGAACAA	67
QY	62	TGATTTTCATTTTACAGCTGGGACAGAGTATTCATTCGCGCCGTTATGACGTGAA	121
DB	68	TGATTTTCATTTTACAGCTGGGACAGAGTATTCATTCGCGCCGTTATGACGTGAA	127
QY	122	ACAGTGAAGCTGACGAGAGGCGCCATCAGACACAGTACCGCTTATCGGTAGGAGAA	181
DB	128	AAATTGATCTTAATTCAGACATGCGCTT---GAATATGTTTGTGAAACAATTTAA	184
QY	182	AACTTGGAATATATTTATTTGCGACCGCATTAAGAAAGCGTTATCTGTGCGCGC	241
DB	185	AACTTGGAATATATTTATTTGCGACCGCATTAAGAAAGCGTTATCTGTGCGCGC	244

[illegible]

Db	1098	GTGGTAAAGCTCACCAATGCGTTTAAACCGCGTACCTAATTTACCAATGATGCGCGCG	781
Qy	722	CGCGCGCCGATGAAACGAAAGGTGTAAACGCCGCTTTCAAAATTTGGCAAAATATGACGA	781
Db	1038	GTTCCTCCATTAATGAAGATGAAATTTAAATGTAATGTGTTAAAGTTCCTAAGATCAGTCAGTA	979
Qy	782	TACCAAAAGGCGCATGTGATTAACGGAAGAGACGAAATTCATTATTAACGAATTAATCTCCG	841
Db	978	TCCCTCAAGGTGCAGTTAT--CGATGCCATTAATAATACATTACACAAATATCAATTAG	922
Qy	842	TGATGTGCACGAAACTGGAAATCTACTATTTCCACCACTATGACAAATCGGCAAAATCCAA	901
Db	921	TGAATGAAAGTAAAGAAAGAAAGTATTAATTAAGCTTACTTTAGCAATCAAAATTTCA	862
Qy	902	AAGTAAATTTATTTTCATGA	920
Db	861	AAATTAATTTACTGAAGA	843

RESULT 3  
 US-10-793-626-4184  
 ; Sequence 4184, Application US/10793626  
 ; Date 1997-04-14, File 4184-1

; GENERAL INFORMATION:

```

APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIORITY FILING DATE: 1999-11-09
PRIORITY FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4184
LENGTH: 4114
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4184

```

```

Query Match      16.0%; Score 159; DB 8; Length 4114;
Best Local Similarity 49.8%; Pred. No. 4e-35;
Matches 458; Conservative 0; Mismatches 455; Indels 6; Gaps 2;

```

```

QY 2 ATATGACACAGCTTACATTTGGAACCTGTCACCGTAACATGATTTAGCAAGACAA 61
DB 1626 ATATGACACAGCTTACATTTGGAACCTGTCACCGTAACATGATTTAGCAAGACAA 1685
QY 62 TGGATTTTGGATTTGACCTGGGGAACAGAGTGAATCTCTAATCCGCGCGTTACAGCTGA 121
DB 1686 TGGATTTTGGATTTGACCTGGGGAACAGAGTGAATCTCTAATCCGCGCGTTACAGCTGA 1745
QY 122 ACAGTGAAGCTGACGAGAGGCGCCATGACACAGTACCGGTTATCGGTATGGGAGAA 181
DB 1746 AATTTGATCTAGATTCAGACATGCTCTT--GAAATGTTGTTGTTGGAACAAATTTAA 1802
QY 182 AACTGGAATATATATTTATTTGCGAGCGCATTAATGAACCGGTTATCTTGGCGCGC 241
DB 1803 AACTGGAATATATATTTATTTGCGAGCGCATTAATGAACCGGTTATCTTGGCGCGC 1862
QY 242 TTTATTTTCCGGCTATGCGAGTACGAAAAAGATACCGGAGATACCGTTCACTTG 301
DB 1863 ATTACTTCACTGGTGAAGCTCATACAGTACCCATTAAGCTTATGTTATTTAATCTAG 1922
QY 302 TCCCGCATGATTTGACATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
DB 1923 CACTGAGAGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 1982
QY 362 AAAAGATGATTTTAAAGATGTTGAGAAATTAATGATCTATTTGATGATGATTTTAC 421
DB 1983 AAAAGATGATTTTAAAGATGTTGAGAAATTAATGATCTATTTGATGATGATTTTAC 2042
QY 422 CGCTTCACTGATATTTGTCAGACCGGACCGGCGGAAACCTGACGATAGAACCAAGACAG 481
DB 2043 CTTTATATTTCACTGATATTTGTCAGACCGGACCGGCGGAAACCTGACGATAGAACCA 2102
QY 482 AGCGCTCAAAGTCTATGATATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
DB 2103 GCTTATATTAAGTAAAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2162
QY 542 GCGATTAACCATCTGACGATATATACGAAATCAACGAAATTAATGAGAGAGAA 601
DB 2163 GCGATTAACCATCTGACGATATATACGAAATCAACGAAATTAATGAGAGAGAA 2222
QY 602 AGATGGCGGATTAAGCCCTTTCTGCTTTGGCCAGGCTTAGGACGTTGCTGCTGCTGCTG 661
DB 2223 TATATGATTAAGTCAAGTATGATCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2282
QY 662 GGGATTAACACCGCTTCCCGGTTGCTGACAGCTGTTATTTGAAAGAACATCTAGAC 721
DB 2283 GTGCTTATATGCTCAACAGATCGTTTATATGCGCTCAATATTAAGACCAACTAGCT 2342
QY 722 CGGCGGCGATGAAGAAAGGATGAACGCGCTTTGAAATTTTGGGAAATATGACGA 781
DB 2343 GTTCCATTAATGAAGTGAAGATTTTATGATGTTTAAAGTTCTAGAAATCACTGACGA 2402

```

```

QY 782 TACCAAGGCGGATGATTAACGGAAGAAAGCAATTTCAATATACCAATATCTCG 841
DB 2403 TCCCTCAAGGTGAGATTTAT--CGATGCCAATTAATATATATATATATATATATAT 2459
QY 842 TGAATGCAACGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 901
DB 2460 TGAATGCAACGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2519
QY 902 AAGTAAATTTATTTCAATGA 920
DB 2520 AAGTAAATTTATTTCAATGA 2538

```

# RESULT 4

```

US-11-074-176-367
Sequence 367, Application US/11074176
Publication No. US20050250135A1
GENERAL INFORMATION:
APPLICANT: Kleenhammer, Todd R.
APPLICANT: Russell, William M.
APPLICANT: Altermann, Eric
APPLICANT: McAniff, Olivia
APPLICANT: Pettit, Andrea Azcarate
TITLE OF INVENTION: Nucleic Acid Sequences Encoding
FILE REFERENCE: 5051-694
CURRENT APPLICATION NUMBER: US/11/074,176
CURRENT FILING DATE: 2005-03-07
PRIORITY FILING DATE: 2004-03-08
PRIORITY FILING DATE: 2004-03-08
NUMBER OF SEQ ID NOS: 381
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 367
LENGTH: 975
TYPE: DNA
ORGANISM: Lactobacillus acidophilus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(975)
FEATURE:
NAME/KEY: misc.feature
LOCATION: (0)...(0)
OTHER INFORMATION: ORF 892; bsha
US-11-074-176-367

```

```

Query Match      9.3%; Score 93; DB 12; Length 975;
Best Local Similarity 46.4%; Pred. No. 2e-16;
Matches 340; Conservative 0; Mismatches 390; Indels 3; Gaps 1;

```

```

QY 197 TATTTGCGAGCGCATTAATGAAGCGGTTTATCTTGTGCGGCGCTTTATTTTCCGGGCT 256
DB 188 TATATTTGCGAGCGCATTAATGAAGCGGTTTATCTTGTGCGGCGCTTTATTTTCCGGGCT 247
QY 257 ATGCGAGTACGAAAAAGATACGGAAGATACCGTTCAATTCGCGATGAGTTTG 316
DB 248 ATGCGAGTACGAAAAAGATACGGAAGATACCGTTCAATTCGCGATGAGTTTG 304
QY 317 TGAATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376
DB 305 TCCCTTGAATTTTGAAGACGTTGATGATGATGATGATGATGATGATGATGATGATGAT 364
QY 377 TAAAGTTTGAAGAAAAATTAATGATCTATTTGATGATGATGATGATGATGATGATGAT 436
DB 365 TAAAGTTTGAAGAAAAATTAATGATCTATTTGATGATGATGATGATGATGATGATGAT 424
QY 437 TGTCAAGCGGACGCGCGGAAACCTGACGATGAACCAAGACAGACGCGCTCAAGTCT 496
DB 425 TGTCAAGCGGACGCGCGGAAACCTGACGATGAACCAAGACAGACGCGCTCAAGTCT 484
QY 497 ATGATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
DB 485 ATGATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544

```

```

Qy 557 TGCGACATATATACAGAAATACAGCCGACATTAAGAGCAAGATGGCGGATTAG 616
Db 545 TAAATTAATTAATGCTGACGTAATCTCCAAAATGCTTAAATAATCTTCAAGTAAGTA 604
Qy 617 CCCTTCTGCTTTTGGCCAAAGGCTTAGAACTGTGTCTGCGCGGGGATTAATACACCGC 676
Db 605 AATAGGCTGGCTACAGCCGCTGATTAAGGCTCACAACCTAACAGGTGAATGATCTG 664
Qy 677 CTCCCGGCTTTGTCAGACTGTTTATTTGAAAAGAACATCTAGAGCCGCGCCATGAAA 736
Db 665 AATACGTTTGTGAGAGTAGCTTTCATTAATTAATTAATGCTTCAATTTCTGAAAACCGAG 724
Qy 737 CGAAAGGTGTAACAGCCGCTTTCAATTTTGGCAATATATGATGATACAAAGGCGGAG 796
Db 725 AAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 784
Qy 797 TGATTAACGGAAGAGACCAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 856
Db 785 ATGAAGTTGTCGAACATCAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 844
Qy 857 CTGGAATCTACTATTTCACCACTATGACATGCGCAATCCAAAAGTAATTTATTTTC 916
Db 845 AAGGATTTTCTACTACACCACTATTCACAAACAAATTAATTAATTAATTAATTAATTAAT 904
Qy 917 ATGAAGCTTGA 929
Db 905 AAGAGATCTAGA 917

```

## RESULT 5

```

US-11-074-176-369
; Sequence 369, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAniff, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Lactobacillus acidophilus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(975)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: ORF 1078; bshb
US-11-074-176-369

```

Query Match 8.4%; Score 83.8; DB 12; Length 975;

Best Local Similarity 45.4%; Pred. No. 8.9e-14;

Matches 341; Conservative 0; Mismatches 407; Indels 3; Gaps 1;

```

Qy 181 AAACCTGAAATATATATTTTCCGACGCGCATTAATGAAGCGTTATCTGTGCGGCG 240
Db 172 AATGATGATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 231
Qy 241 CTTTATTTTCCGGGCTATGCGGAGTACGAAAAACGATCGGAGAGATCCGTTCACTT 300
Db 232 TTAATTTTCAAGG--TCTTAATCACTTCTTCTTAATTAATGAAGTAAAGATATTT 288

```

```

Qy 301 GTCCCGATGAGTTTGATGACATGGGTGCTGCTGATCTGATCTGTTTGAAGAGTAATA 360
Db 289 GCTTCTTTGATTAATTAATGACATCTTATTAAGTAATTTGTAATAATTAATGACATGTTAAA 348
Qy 361 GAAAGATTCGATCTTTAATGATTTGAGAAAAAATTAATTAATTAATTAATTAATTAATTAAT 420
Db 349 GAAATCTTAATTAATGCAAAATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 408
Qy 421 CCGCTTACGATATGTCAGACCGGACGGGCGAAACCTGACGATAGAACCAAGAGCA 480
Db 409 GATTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 468
Qy 481 GACGCGCTCAAGATCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
Db 469 TGAGTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 528
Qy 541 TGGCATGTAACCAATCTGACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Db 529 GATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 588
Qy 601 GAGATGGCGGATTAAGCCCTTCTGCTTTGGCCAGGCTTAGGAGCTGTGCTGCGG 660
Db 589 TTGTTCTTAATGTTGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 648
Qy 661 GGGGATTAATACACCGCTTCCCGGTTTGTGACAGCTGTTTATTTGAAGAATCATCTAGAG 720
Db 649 GGTGGAATGATTTTACTCTGATTTGTTAAGTACTTTTGTCTGGACACATCTCA 708
Qy 721 CCGGCGGCGCATGAACGAAAGGTGTAACAGCCGCTTTCAATTTTGGCAATTAATGACG 780
Db 709 CAAGAAAAAATGAAGGAAATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 768
Qy 781 ATACCAAGGCGGAGGATTAAGGAAAGAGACAAATTCATTAATGCAATTAATTAATTAATTAAT 840
Db 769 CAACCTGATGTTTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 828
Qy 841 GTGATGTGCAACGAAGCTGAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
Db 829 TGTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 888
Qy 901 AAGTAATTTATTTCAATGAAGCTTGACT 931
Db 889 GCAGTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 919

```

## RESULT 6

```

US-10-750-185-59863/c
; Sequence 59863, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 59863
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-59863

```

Query Match 3.9%; Score 38.8; DB 8; Length 720;

```
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```

```

; NAME/KEY: misc_feature
; LOCATION: (1)..(1272)
; OTHER INFORMATION: Ceres Seq. ID no. 4809357
US-11-0936-568A-29311

```

Query Match	3.6%;	Score 36.2;	DB 9;	Length 1272;
Best Local Similarity	57.5%;	Pred. No. 5.2;		

149 AGACACAGTACCGCTTATCGGTATGGGAGAAAATTGGAATATATATTTCGCACG 208

Db 479 GCTTCAATGGTACCGATGATCTCTCCGGCCTTATCTATCTTACGTTGTG 531

US-09-925-065A-392869/c  
; Sequence 392869, Application US/09925065A  
; Publication No. US20040181048A1

```

; GENERAL INFORMATION:
;
; APPLICANT: Wang, David G.
;
; TITLE OF INVENTION: Identification and Mapping of Single
;
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

```

1 CURRENT APPLICATION NUMBER: US 08/09,925, 0654  
2  
3 CURRENT FILING DATE: 2001-08-08  
4  
5 PRIOR APPLICATION NUMBER: US 60/243, 036  
6  
7 PRIOR FILING DATE: 2000-10-24  
8  
9 PRIOR APPLICATION NUMBER: US 60/252, 147  
10  
11 FILING DATE: 2000-11-20  
12  
13 PRIOR APPLICATION NUMBER: US 60/250, 092

PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846

```

; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 392869

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-392869

```

Query Match	3.6%;	Score	35.8;	DB	6;	Length	583;
Best Local Similarity	49.7%;	Pred. No.	4	6;			
Matches	88;	Conservative	1;	Mismatches	88;	Indels	0;
						Gaps	0

QY 802 ACGGAAGAAGACGAAATCATATTACGCATATATCTCCGTATGTGCAACGAACCTGGA 861

Db 341 ATGGAAGAGCTTAATGATATTATTCACAACTAAAAAGAACATCTGAAAAGACTGCATA 282

862 AACCTACTATTTCACCCATATGACATTCGCAATTCGCAAAATCCAAAAGTAATTATTTCAATGAA 921  
 281 CTGATATGATTCACACTATATGACATTCGCAAAAACAAAAACCATGTAAATAGTAAAAA 222

922 GACCTTACACGTTGAGACCCAAAGAGTATTCAT 9/8  
 ||| : ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 221 GATCAGTGGTTCGCGGGGAAAACGGGGGATGAA CAGTGGCGCACAGAGGATTCCTT 165

RESULT 10  
US-09-925-065A-654182  
Semence 654182 Application US/09925065A

APPLICANT: Wang, David G

;; TITLE OF INVENTION: Identification and Mapping of Single  
;; FILE REFERENCE: 108827.135  
;; CURRENT APPLICATION NUMBER: US/09/925,065A  
;; PRIOR FILING DATE: 2001-08-08  
;; PRIOR APPLICATION NUMBER: US 60/243,096  
;; PRIOR FILING DATE: 2000-10-24  
;; PRIOR APPLICATION NUMBER: US 60/252,147  
;; PRIOR FILING DATE: 2000-11-20  
;; PRIOR APPLICATION NUMBER: US 60/250,092  
;; PRIOR FILING DATE: 2000-11-30  
;; PRIOR APPLICATION NUMBER: US 60/261,766  
;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/289,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 654182  
;; LENGTH: 419  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-654182

Query Match 3.6%; Score 35.4; DB 6; Length 419;  
Best Local Similarity 57.8%; Pred. No. 5.1;  
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 333 AGCTGTCAGCTTGGAGACGTAAAGAAAGATTGATCTTAAAGATTGAGAA 392  
Db 244 AGATTAAACAGATTAAAGAAACAGATTTTAAAGATTAATTAACCACTTTATGAA 303  
Qy 393 AAAATTAGATCTATTGGATACAGTTTACCGCTTCAGTGAATTTGCA 441  
Db 304 TAAATTGATGACAGATTATTTTCTTCGATCTGTAATTTGTA 352

RESULT 11  
US-09-925-065A-654183  
;; Sequence 654183, Application US/09925065A  
;; Publication No. US20040181048A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single  
;; FILE REFERENCE: 108827.135  
;; CURRENT APPLICATION NUMBER: US/09/925,065A  
;; PRIOR FILING DATE: 2001-08-08  
;; PRIOR APPLICATION NUMBER: US 60/243,096  
;; PRIOR FILING DATE: 2000-10-24  
;; PRIOR APPLICATION NUMBER: US 60/252,147  
;; PRIOR FILING DATE: 2000-11-20  
;; PRIOR APPLICATION NUMBER: US 60/250,092  
;; PRIOR FILING DATE: 2000-11-30  
;; PRIOR APPLICATION NUMBER: US 60/261,766  
;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/289,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 654183  
;; LENGTH: 419  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-654183

Query Match 3.6%; Score 35.4; DB 6; Length 419;  
Best Local Similarity 57.8%; Pred. No. 5.1;  
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 333 AGCTGTCAGCTTGGAGACGTAAAGAAAGATTGATCTTAAAGATTGAGAA 392  
Db 244 AGATTAAACAGATTAAAGAAACAGATTTTAAAGATTAATTAACCACTTTATGAA 303

Qy 393 AAAATTAGATCTATTGGATACAGTTTACCGCTTCAGTGAATTTGCA 441  
Db 304 TAAATTGATGACAGATTATTTTCTTCGATCTGTAATTTGTA 352

RESULT 12  
US-10-932-182A-2969  
;; Sequence 2969, Application US/10932182A  
;; Publication No. US20060046253A1  
;; GENERAL INFORMATION:  
;; APPLICANT: NAKAO, YOSHIHIRO  
;; APPLICANT: KODAMA, YUKIKO  
;; APPLICANT: FUJIMURA, TOMOKO  
;; APPLICANT: ASHIKARI, TOSHIHIKO  
;; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
;; FILE REFERENCE: 030685-043  
;; CURRENT APPLICATION NUMBER: US/10/932,182A  
;; PRIOR FILING DATE: 2004-09-02  
;; NUMBER OF SEQ ID NOS: 197023  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO: 2969  
;; LENGTH: 5901  
;; TYPE: DNA  
;; ORGANISM: Saccharomyces pastorianus  
US-10-932-182A-2969

Query Match 3.6%; Score 35.4; DB 7; Length 5901;  
Best Local Similarity 50.9%; Pred. No. 19;  
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 693 AGCTGTTATTGGAAGACATCTAGAGCGGCGCGGATGAAACGAAAGGTGAACGC 752  
Db 894 AGCTTATACACGATGAAATATGTTGATTTGCCCTACAGTGCAGGTAAACAGA 953  
Qy 753 CGCTTTCAATTTTGGCAATATGACATACCAAGGCGGAGTATTAACGAAGAGA 812  
Db 954 CGTTCACACTTACGATTTAAACACGTTAAACAGTTTCTCCGTATCAACGAGAAAA 1013  
Qy 813 CGAATTCAATTATACGCAATATCTTCCGTGATGCAAGAAAC 857  
Db 1014 CGAATTGATATCAATACGATAGTTTCAAGGTTGTCTAGTAC 1058

RESULT 13  
US-10-932-182A-2969  
;; Sequence 2969, Application US/10932182A  
;; Publication No. US20060046253A1  
;; GENERAL INFORMATION:  
;; APPLICANT: NAKAO, YOSHIHIRO  
;; APPLICANT: KODAMA, YUKIKO  
;; APPLICANT: FUJIMURA, TOMOKO  
;; APPLICANT: ASHIKARI, TOSHIHIKO  
;; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
;; FILE REFERENCE: 030685-043  
;; CURRENT APPLICATION NUMBER: US/10/932,182A  
;; PRIOR FILING DATE: 2004-09-02  
;; NUMBER OF SEQ ID NOS: 197023  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO: 2969  
;; LENGTH: 5901  
;; TYPE: DNA  
;; ORGANISM: Saccharomyces pastorianus  
US-10-932-182A-2969

Query Match 3.6%; Score 35.4; DB 7; Length 5901;  
Best Local Similarity 50.9%; Pred. No. 19;  
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 693 AGCTGTTATTGGAAGACATCTAGAGCGGCGGAGTAAACGAAAGGTGAACGC 752  
Db 894 AGCTTATACACGATGAAATATGTTGATTTGCCGTACAGGTGACGTTAAACAGA 953



QY 753 CGCTTTCAATTGGCAATATGACGATACCAAGGGCGAGTGATAACGAGAGAA 812  
DB 954 CGTTGACACTTACGATATTTAAACAGGTAAACAGTTCTCCGTGATCAACGAGAAA 1013  
QY 813 CGAAATTCATTATACGCAATATCTCCGATGTGCAAGAAC 857  
DB 1014 CGAAATTCATTATACATATACGATGTTCAAGTTGTCAGTAGC 1058

## RESULT 14

US-11-072-175-102  
; Sequence 102, Application US/11072175  
; Publication No. US20060299441  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT  
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS  
; FILE REFERENCE: D0273A CIP  
; CURRENT APPLICATION NUMBER: US/11/072,175  
; PRIOR FILING DATE: 2005-03-05  
; PRIOR APPLICATION NUMBER: US 60/406,385  
; PRIOR FILING DATE: 2002-08-27  
; PRIOR APPLICATION NUMBER: US 10/648,593  
; PRIOR FILING DATE: 2003-08-26  
; NUMBER OF SEQ ID NOS: 571  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 102  
; LENGTH: 961  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-072-175-102

Query Match 3.5%; Score 35; DB 9; Length 961;  
Best Local Similarity 47.1%; Pred. No. 10; Mismatches 120; Indels 0; Gaps 0;  
Matches 107; Conservative 0;

QY 293 TTCACATGTCGCCGATGAGTTGTGACATGGGTGTCAGTCTGCACTTTTGAAG 352  
DB 466 TTAAACACAGTCATGAAAGTTGGCTTCGTTGATGTCTGTTGGCGCTGGAAC 525  
QY 353 ACGTAAAGAAAAGATTCATCTTTAAGCATTTAGAGAAAAATTGATCTATTGATA 412  
DB 526 AAGGACACCACTATGAGATGTTCTGTGTTCAGTGAAAAATTCATTATGAGA 585  
QY 413 CAGTTTACCGCTTCACTGATATGTGACAGCCGAGCGGCGAAACCTGAGATGAAC 472  
DB 586 CATTTTCTTTTTCACAGTATTTTTCACACTGTATCATGAAACTACCTTTAGAA 645  
QY 473 CAAGACAGACGGCTCAAGTCTATGATATCACTGATGATGATG 519  
DB 646 AATAAATTATACCTGGCAAAAATAAGAGAAACGAAACTGACGTTATG 692

RESULT 15  
US-09-925-065A-392868/c  
; Sequence 392868, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2003-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 392868  
; LENGTH: 583  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-392868

Query Match 3.5%; Score 34.6; DB 6; Length 583;  
Best Local Similarity 49.7%; Pred. No. 10; Mismatches 89; Indels 0; Gaps 0;  
Matches 88; Conservative 0;

QY 802 ACGAAGAGAGAAATTCATTATACGCAATATCTCCGTGATGCAACGAAACTGA 861  
DB 341 ATGGAAGAGCTTAATGATATTTACAGTAAGAAAGACATCTGAAAGACTGATA 282  
QY 862 AACTACTATTTCACCACTATGACATCGCAATCCAAAAGTAATTTATTTATGAA 921  
DB 281 CTGATATGATTCACACTATATGACATTCAGAAAACAAAACCATGTAAATGTAAAA 222  
QY 922 GACTTGACTGTTTGAAGCTTAAAGTTTCCGCTAAAGCAGAGAGATTTCT 978  
DB 221 GATCAGTGTGTCACAGGGGAAAAACGGGGGTGAACAGGTGCGCACAGAGATTTCT 165

Search completed: March 18, 2006, 12:06:10  
Job time : 453 sec